

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2002, 23:33:49 ; Search time 1145.36 Seconds

(without alignments)
23,980 Million cell updates/sec

Title: US-09-672-126-7

Perfect score: 20

Sequence: 1 ggggggacgacgcgcggggg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: N.GeneSeq_032802.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	20	100.0	20	AAF98737
2	20	100.0	20	AAF98852
3	20	100.0	21	AAF98745
4	20	100.0	21	AAF98746
5	20	100.0	21	AAF99791
6	20	100.0	21	AAF99792
7	19	95.0	19	AAF99754
8	19	95.0	20	AAF98765
9	19	95.0	22	AAF99869

ALIGNMENTS

10	18.4	92.0	20	22	AAF98743	Human IFN-alpha im
11	18.4	92.0	20	22	AAF98744	Human IFN-alpha im
12	18.4	92.0	20	22	AAF99789	Immunostimulatory
13	18.4	92.0	20	22	AAF99790	Immunostimulatory
14	18	90.0	19	22	AAF98762	Human IFN-alpha im
15	18	90.0	19	22	AAF98865	Immunostimulatory
16	16.8	84.0	20	22	AAF98742	Human IFN-alpha im
17	16.8	84.0	20	22	AAF99786	Immunostimulatory
18	16.8	84.0	699	21	AAF13496	Aspergillus oryzae
19	16.4	82.0	20	22	AAF99850	Immunostimulatory
20	15.4	77.0	4011	15	AAQ70362	Varicella zoster v
21	15.4	77.0	4011	15	AAQ69881	DNA encoding VZV I
22	15.4	77.0	4226	21	AAA93780	Chicken-pox virus
23	15.4	77.0	124884	22	AAH74201	Nucleotide sequenc
24	15.4	77.0	124884	22	AAH74201	Nucleotide sequenc
25	15.4	77.0	125157	22	AAH74202	Nucleotide sequenc
26	15.4	77.0	125157	22	AAH74202	Nucleotide sequenc
27	15.2	76.0	125157	20	AAF98752	Human IFN-alpha im
28	15.2	76.0	20	22	AAF98761	Human IFN-alpha im
29	15.2	76.0	20	22	AAF98876	Immunostimulatory
30	15.2	76.0	20	22	AAF99835	Immunostimulatory
31	15.2	76.0	20	22	AAF99848	Immunostimulatory
32	15.2	76.0	20	22	AAF99849	Immunostimulatory
33	15.2	76.0	882	19	AAV64512	M. tuberculosis im
34	15.2	76.0	882	19	AAV64512	M. tuberculosis im
35	15.2	76.0	882	20	AAZ19313	M. tuberculosis im
36	15.2	76.0	882	20	AAZ19101	M. tuberculosis re
37	15.2	76.0	1299	23	AAZ19101	DNA encoding novel
38	15.2	76.0	1436	21	AAZ19101	Arabidopsis thalia
39	15.2	76.0	1559	22	AAZ19101	Human immune/haema
40	15.2	76.0	2453	23	ABL02487	Drosophila melanog
41	15.2	76.0	3648	21	AAZ19101	Human ORFX ORF1415
42	15.2	76.0	4444	22	AAI59840	Human polynucleoti
43	15.2	76.0	4446	22	AAI59840	Human polynucleoti
44	15.2	76.0	4867	23	ABL04655	Drosophila melanog
45	15.2	76.0	7529	23	ABL02486	Drosophila melanog

RESULT 1
AAF98737
ID AAF98737 standard; DNA; 20 BP.

AAF98737;

11-JUN-2001 (first entry)

Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 7.

Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha; viral infection; phosphorothioate backbone; palindrome; cancer; ds.

Synthetic.

key Location/Qualifiers

modified_base 1..2

modified_base 15..19

modified_base 15..19

modified_base 15..19

modified_base 15..19

modified_base 15..19

modified_base 15..19

modified_base 15..19

modified_base 15..19

modified_base 15..19

XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.
XX
XX Hartmann G, Bratzler RL, Krieg A;
PI WPI; 2001-290487/30.
XX
XX Improving the efficacy of treatments involving the administration of
PT Interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
XX Claim 201; Page 103; 168pp; English.
XX
XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX
XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;
SQ

Query Match 100.0%; Score 20; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 99999acgacgtcgcg9999 20
|||||
Db 1 99999acgacgtcgcg9999 20

RESULT 2
AAF98852
ID AAF98852 standard; DNA; 20 BP.
XX
XX AAF98852;
XX
DT 11-JUN-2001 (first entry)
XX
XX Poly-G immunostimulatory nucleic acid SEQ ID NO: 133.
XX
XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KM viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
XX Synthetic.
OS
XX WO200122990-A2.
PN
XX
XX 05-APR-2001.
PD
XX
XX 27-SEP-2000; 2000WO-US26527.
PF
XX
XX 27-SEP-1999; 99US-0156147.
PR
XX
XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.
XX
XX Hartmann G, Bratzler RL, Krieg A;
PI
XX
XX WPI; 2001-290487/30.
DR
XX
XX Improving the efficacy of treatments involving the administration of
PT Interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
XX Disclosure; Page 24; 168pp; English.
PS
XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering

CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX
XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;
SQ

Query Match 100.0%; Score 20; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 99999acgacgtcgcg9999 20
|||||
Db 1 99999acgacgtcgcg9999 20

RESULT 3
AAF98745
ID AAF98745 standard; DNA; 21 BP.
XX
XX AAF98745;
XX
XX 11-JUN-2001 (first entry)
DT
XX
XX Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 15.
DE
XX
XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KM viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH modified_base 1..2
FT /*tag= a
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
FT modified_base 16..20
FT /*tag= b
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
XX
XX WO200122990-A2.
PN
XX
XX 05-APR-2001.
PD
XX
XX 27-SEP-2000; 2000WO-US26527.
PF
XX
XX 27-SEP-1999; 99US-0156147.
PR
XX
XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.
XX
XX Hartmann G, Bratzler RL, Krieg A;
PI
XX
XX WPI; 2001-290487/30.
DR
XX
XX Improving the efficacy of treatments involving the administration of
PT Interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
XX Claim 201; Page 103; 168pp; English.
PS
XX
XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.


```
RESULT 6
AAF99792 standard; DNA: 21 BP.
XX
AC AAF99792;
XX
DT 12-JUN-2001 (first entry)
XX
DE Immunostimulatory nucleic acid #908.
XX
KM Vaccine: cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KM immunostimulatory; tumour; viral infection; bacterial infection;
KM fungal infection; parasitic infection; cancer; asthma;
KM infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS Synthetic.
XX
PN WO200122972-A2.
XX
PD 05-APR-2001.
XX
PF 25-SEP-2000; 2000WO-US26383.
XX
PR 25-SEP-1999; 99US-0156113.
PR 27-SEP-1999; 99US-0156135.
PR 23-AUG-2000; 2000US-0227436.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX
DR WPI; 2001-273485/28.
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 58; 338pp; English.
XX
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-potent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SQ Sequence 21 BP; 2 A; 3 C; 14 G; 2 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggggagcagtcgcggggg 20
   ||||||||||||||||
DB 1 gggggagcagtcgcggggg 20

RESULT 7
AAF99754 standard; DNA: 19 BP.
XX
AC AAF99754;
XX
DT 12-JUN-2001 (first entry)
XX
```

```
XX
DE Immunostimulatory nucleic acid #870.
XX
KM Vaccine: cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KM immunostimulatory; tumour; viral infection; bacterial infection;
KM fungal infection; parasitic infection; cancer; asthma;
KM infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS Synthetic.
XX
PN WO200122972-A2.
XX
PD 05-APR-2001.
XX
PF 25-SEP-2000; 2000WO-US26383.
XX
PR 25-SEP-1999; 99US-0156113.
PR 27-SEP-1999; 99US-0156135.
PR 23-AUG-2000; 2000US-0227436.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX
DR WPI; 2001-273485/28.
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 57; 338pp; English.
XX
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-potent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SQ Sequence 19 BP; 2 A; 3 C; 12 G; 2 T; 0 other;

Query Match 95.0%; Score 19; DB 22; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggggagcagtcgcggggg 19
   ||||||||||||||||
DB 1 gggggagcagtcgcggggg 19

RESULT 8
AAF98765 standard; DNA: 20 BP.
XX
AC AAF98765;
XX
DT 11-JUN-2001 (first entry)
XX
DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 35.
XX
KM Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KM viral infection; phosphorothioate backbone; palindromic; cancer; ds.
XX
OS Synthetic.
```

```

XX 25-SEP-2000; 2000OWO-US26383.
PF 25-SEP-1999; 99US-0156113.
XX 27-SEP-1999; 99US-0156135.
PR 23-AUG-2000; 2000US-0227436.
XX (IOWA ) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
XX Krieg AM, Schetter C, Vollmer J;
XX WPI; 2001-273485/28.
DR
XX
XX Vaccinating against tumors, infectious diseases, allergies and asthma
PT using Immunostimulatory Py-rich and Tg nucleic acids -
XX
XX Claim 101; Page 59; 338BP; English.
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC staphylococcus), campylobacter, clostridium, Escherichia coli and/or
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Tn2 to a Tn1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;
XX
XX
XX Query Match 95.0%; Score 19; DB 22; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 7.7;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX 2 ggggacgacgtcggggg 20
XX |||||||
XX 1 ggggacgacgtcggggg 19
XX
XX
XX RESULT 10
XX AAF98743
XX ID AAF98743 standard; DNA; 20 BP.
XX
XX AAF98743:
XX
XX 11-JUN-2001 (first entry)
XX
XX Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 13.
XX
XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
XX viral infection; phosphorothioate backbone; palindromic; cancer; ds.
XX
XX Synthetic.
XX
XX
XX Key Location/Qualifiers
XX modified_base 1..2
XX /*tag= a
XX /mod_base= "OTHER"
XX /note= "phosphorothioate linkage"
XX modified_base 16..19
XX /*tag= b
XX /mod_base= "OTHER"
XX /note= "phosphorothioate linkage"
XX
XX WO200122990-A2.
XX

```

PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000MO-US26527.
XX
PR 27-SEP-1999; 99US-0156147.
XX
PA (COLE-) COLEY PHARM GROUP INC.
XX (IOWA) UNIV IOWA RES FOUND.
PI Hartmann G, Bratzler RL, Krieg A;
XX WPI; 2001-290487/30.
DR
XX
PT Improving the efficacy of treatments involving the administration of
PT Interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
PS Claim 201; Page 103; 168pp; English.
XX
XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX
SQ Sequence 20 BP; 2 A; 2 C; 13 G; 3 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 1 9999gacgacgtcgcggggg 20
Db 1 9999gacgacgtcgcggggg 20

RESULT 11
AA98744
ID AAF98744 standard; DNA: 20 BP.
XX
AC AAF98744;
XX
DT 11-JUN-2001 (first entry)
XX
DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 14.
XX
XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
XX Synthetic.
OS
FH Location/Qualifiers
FT modified_base 1..2
FT /*tag- a
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
FT 16..19
FT /*tag- b
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
XX
XX WO200122990-A2.
XX
XX 05-APR-2001.
XX
XX 27-SEP-2000; 2000MO-US26527.
XX
XX 27-SEP-1999; 99US-0156147.
XX

PA (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.
XX
PI Hartmann G, Bratzler RL, Krieg A;
XX WPI; 2001-290487/30.
DR
XX
XX Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
PS Claim 201; Page 103; 168pp; English.
XX
XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX
SQ Sequence 20 BP; 3 A; 3 C; 12 G; 2 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 1 9999gacgacgtcgcggggg 20
Db 1 9999gacgacgtcgcggggg 20

RESULT 12
AA99789
ID AAF99789 standard; DNA: 20 BP.
XX
AC AAF99789;
XX
DT 12-JUN-2001 (first entry)
XX
DE Immunostimulatory nucleic acid #905.
XX
XX Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
XX Synthetic.
OS
PN WO200122972-A2.
XX
PD 05-APR-2001.
XX
XX 25-SEP-2000; 2000MO-US26383.
XX
XX 25-SEP-1999; 99US-0156113.
XX 27-SEP-1999; 99US-0156135.
PR 23-AUG-2000; 2000US-0227436.
XX
XX (IOWA) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
XX Krieg AM, Schetter C, Vollmer J;
XX WPI; 2001-273485/28.
XX
XX Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 58; 338pp; English.
XX

CC The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 CC
 SQ Sequence 20 BP; 2 A; 2 C; 13 G; 3 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;
 Best Local Similarity 95.0%; Pred. No. 15;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 99999999999999999999 20
 Db 1 99999999999999999999 20

RESULT 13

AAF99790 ID AAF99790 standard; DNA; 20 BP.

AC AAF99790;

DF 12-JUN-2001 (first entry)

DE Immunostimulatory nucleic acid #906.

KM Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 immunostimulatory; tumour; viral infection; bacterial infection;

KW fungal infection; parasitic infection; cancer; asthma;

KM infectious disease; allergy; immune deficiency; phosphorothioate; ss.

OS Synthetic.

PN WO200122972-A2.

PD 05-APR-2001.

PF 25-SEP-2000; 2000WO-US26383.

PR 25-SEP-1999; 99US-0156113.

PR 27-SEP-1999; 99US-0156135.

PR 23-AUG-2000; 2000US-0227436.

PA (IOWA) UNIV IOWA RES FOUND.

PA (COLE-) COLEY PHARM GMBH.

PI Krieg AM, Schetter C, Volmer J;

DR WPI; 2001-273485/28.

PT Vaccinating against tumors, infectious diseases, allergies and asthma

PS using immunostimulatory Py-rich and TG nucleic acids -

PS Claim 101; Page 58; 338pp; English.

XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae

CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 CC
 SQ Sequence 20 BP; 3 A; 3 C; 12 G; 2 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;
 Best Local Similarity 95.0%; Pred. No. 15;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 99999999999999999999 20
 Db 1 99999999999999999999 20

RESULT 14

AAF98762 ID AAF98762 standard; DNA; 19 BP.

AC AAF98762;

DF 11-JUN-2001 (first entry)

DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 32.

KM Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
 viral infection; phosphorothioate backbone; palindrome; cancer; ds.

KW Synthetic.

OS Synthetic.

PN WO200122990-A2.

PD 05-APR-2001.

PF 27-SEP-2000; 2000WO-US26527.

PR 27-SEP-1999; 99US-0156147.

PR (COLE-) COLEY PHARM GROUP INC.

PA (IOWA) UNIV IOWA RES FOUND.

PA Hartmann G, Bratzler RL, Krieg A;

PI WPI; 2001-290487/30.

DR Improving the efficacy of treatments involving the administration of

DR interferon-alpha by co-administering an isolated immunostimulatory

DR nucleic acid -

DR Claim 201; Page 103; 168pp; English.

XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is

CC an example of an immunostimulatory oligonucleotide.
XX
SQ Sequence 19 BP; 2 A; 3 C; 12 G; 2 T; 0 other;

Query Match 90.0%; Score 18; DB 22; Length 19;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 gggacgacgcgcggggg 20
|||||
DB 1 gggacgacgcgcggggg 18

RESULT 15

AAF99865
ID AAF99865 standard; DNA; 19 BP.

AC AAF99865;

DT 12-JUN-2001 (first entry)

DE Immunostimulatory nucleic acid #981.

KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
immunostimulatory; tumour; viral infection; bacterial infection;
fungal infection; parasitic infection; cancer; asthma;
infectious disease; allergy; immune deficiency; phosphorothioate; ss.

OS Synthetic.

PN WO200122972-A2.

PD 05-APR-2001.

PF 25-SEP-2000; 2000WO-0526383.

PR 25-SEP-1999; 99US-0156113.

PR 27-SEP-1999; 99US-0156135.

PR 23-AUG-2000; 2000US-0227436.

PA (IOWA) UNIV IOWA RES FOUND.

PA (COLE-) COLEY PHARM GMBH.

PI Krieg AM, Schetter C, Vollmer J;

DR WPI; 2001-273485/28.

PT Vaccinating against tumors, infectious diseases, allergies and asthma

PT using immunostimulatory Py-rich and TG nucleic acids -

PS Claim 101; Page 59; 338pp; English.

XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumor antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SQ Sequence 19 BP; 2 A; 3 C; 12 G; 2 T; 0 other;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 gggacgacgcgcggggg 20
|||||
DB 1 gggacgacgcgcggggg 18

Search completed: August 10, 2002, 03:21:45
Job time: 13676 sec

Query Match 90.0%; Score 18; DB 22; Length 19;
Best Local Similarity 100.0%; Pred. No. 22;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:57:36 ; Search time 2778.35 Seconds
(without alignments)
165,704 Million cell updates/sec

Title: US-09-672-126-9

Sequence: 1 99999acgacatcgcgcg9999 22

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

1: gb_da:*
2: gb_hg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_other:*
33: em_hgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

1	22	100.0	22	6	AX104796	AX104796 Sequence
2	22	100.0	22	6	AX105111	AX105111 Sequence
3	18.8	85.5	22	6	AX104797	AX104797 Sequence
4	18.8	85.5	22	6	AX104798	AX104798 Sequence
5	18.8	85.5	22	6	AX105112	AX105112 Sequence
6	18.8	85.5	22	6	AX105113	AX105113 Sequence
7	17.8	80.9	1865	9	HSMB00028	AL050269 Homo sapi
8	17.8	80.9	1865	9	BC004195	BC004195 Homo sapi
9	17.8	80.9	1952	9	BC004225	BC004225 Homo sapi
10	17.8	80.9	133406	3	AF321227	AF321227 Tribolium
11	17.8	80.9	166044	2	AC016888	AC016888 Homo sapi
12	17.8	80.9	179594	2	AC068874	AC068874 Homo sapi
13	17.8	80.9	310050	1	RME603642	AL603642 Rhizobium
14	17.4	79.1	8718	6	AX346175	AX346175 Sequence
15	17.2	78.2	954	8	OSU11773	U31773 Oryza sativ
16	17.2	78.2	48997	2	AC055728	AC055728 Homo sapi
17	17.2	78.2	110000	2	LMFCH31_18	Continuation (19 o
18	17.2	78.2	127781	2	AC097033	AC097033 Rattus no
19	17.2	78.2	129732	2	AC105356	AC105356 Rattus no
20	17.2	78.2	134682	2	AC092553	AC092553 Oryza sat
21	17.2	78.2	138701	2	AP004187	AP004187 Oryza sat
22	17.2	78.2	141869	2	AC097688	AC097688 Rattus no
23	17.2	78.2	142885	8	AP004127	AP004127 Oryza sat
24	17.2	78.2	147207	2	AC078890	AC078890 Oryza sat
25	17.2	78.2	151359	2	AC025098	AC025098 Oryza sat
26	17.2	78.2	168133	2	AC093622	AC093622 Homo sapi
27	16.8	76.4	5799	8	TVDNALP61	X75655 T. versicolo
28	16.8	76.4	157918	2	AC024252	AC024252 Homo sapi
29	16.8	76.4	166110	2	AC098755	AC098755 Rattus no
30	16.8	76.4	167603	2	AC099471	AC099471 Rattus no
31	16.8	76.4	173658	9	AC073046	AC073046 Homo sapi
32	16.4	74.5	4417	1	AH7276532	AJ276532 Aequonass
33	16.4	74.5	10037	1	AE005939	AE005939 Caulobact
34	16.4	74.5	211009	8	AF326781	AF326781 Trilicium
35	16.2	73.6	21	6	AX104887	AX104887 Sequence
36	16.2	73.6	21	6	AX105139	AX105139 Sequence
37	16.2	73.6	494	3	AF436372	AF436372 Catocala
38	16.2	73.6	1097	3	AF436621	AF436621 Hydropsyc
39	16.2	73.6	1098	3	AF436629	AF436629 Wormaldia
40	16.2	73.6	1227	3	AF056098	AF056098 Colpoda 1
41	16.2	73.6	1419	6	AX078533	AX078533 Sequence
42	16.2	73.6	2148	6	A73577	A73577 Sequence 1
43	16.2	73.6	2448	8	AF326116	AF326116 Agastache
44	16.2	73.6	2642	8	AK001644	AK001644 Homo sapi
45	16.2	73.6	2654	1	STYFLH1J	M62408 Salmonella

ALIGNMENTS

RESULT 1
AX104796
LOCUS AX104796 22 bp
DEFINITION Sequence 988 from Patent WO0122972.
ACCESSION AX104796
VERSION AX104796.1 GI-13920993
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS Krieg, A.M., Schetter, C. and Vollmer, J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 988 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)

FEATURES
source location/Qualifiers
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 3 a 3 c 13 g 3 t
ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gggggacgatactgcg999g 22
|||||
Db 1 GGGGACGATATCGTCGGGGG 22

RESULT 2
AX105111
LOCUS AX105111 22 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 9 from Patent WO0122990.
ACCESSION AX105111
VERSION AX105111.1 GI:13921261
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 9 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source location/Qualifiers
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"

misc_feature 1..22
/note="Backbone has phosphorothioate linkages."
misc_feature 3..16
/note="Backbone has phosphodiester linkages."
misc_feature 17..21
/note="Backbone has phosphorothioate linkages."
misc_feature 22
/note="Backbone has phosphodiester linkages."
BASE COUNT 3 a 3 c 13 g 3 t
ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggggacgatactgcg999g 22
|||||
Db 1 GGGGACGATATCGTCGGGGG 22

RESULT 3
AX104797
LOCUS AX104797 22 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 989 from Patent WO0122972.
ACCESSION AX104797
VERSION AX104797.1 GI:13920994
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 989 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES
source location/Qualifiers
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 2 a 4 c 14 g 2 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 6; Length 22;
Best Local Similarity 90.9%; Pred. No. 8.7e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgatactgcg999g 22
|||||
Db 1 GGGGACGACGCTCGTCGGGGG 22

RESULT 4
AX104798
LOCUS AX104798 22 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 990 from Patent WO0122972.
ACCESSION AX104798
VERSION AX104798.1 GI:13920995
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 990 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES
source location/Qualifiers
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 2 a 4 c 14 g 2 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 6; Length 22;
Best Local Similarity 90.9%; Pred. No. 8.7e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgatactgcg999g 22
|||||
Db 1 GGGGACGACGCTCGTCGGGGG 22

RESULT 5
AX105112
LOCUS AX105112 22 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 10 from Patent WO0122990.
ACCESSION AX105112
VERSION AX105112.1 GI:13921262
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 10 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source location/Qualifiers
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"

misc_feature 1..22
/note="Backbone has phosphorothioate linkages."
misc_feature 3..16
/note="Backbone has phosphodiester linkages."
misc_feature 17..21

```

misc_feature      22 /note="Backbone has phosphorothioate linkages."
BASE COUNT      2 a      4 c      14 g      2 t
ORIGIN
Query Match      85.5%; Score 18.8; DB 6; Length 22;
Best Local Similarity 90.9%; Pred. No. 8.7e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 9999gacgatatcgtcggggg 22
    |||||||  |||||||
Db 1 GGGGGACGACGCTGCTCGGGGG 22

RESULT 6
AX105113      22 bp      DNA      linear      PAT 30-APR-2001
LOCUS      Sequence 11 from Patent WO0122990.
DEFINITION      AX105113
ACCESSION      AX105113
VERSION      AX105113.1 GI:13921263
KEYWORDS
SOURCE      synthetic construct.
ORGANISM      synthetic construct.
REFERENCE      1 (bases 1 to 22)
AUTHORS      Hartmann,G.D., Bratzler,R.L. and Kriegl,A.U.
TITLE      Methods related to immunostimulatory nucleic acid-induced
            interferon
JOURNAL      Patent: WO 0122990-A 11 05-APR-2001;
            Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
            FOUNDATION (US)
FEATURES
    source      location/Qualifiers
                1..22
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="Synthetic Oligonucleotide"
                misc_feature      1..2
                /note="Backbone has phosphorothioate linkages."
                misc_feature      3..16
                /note="Backbone has phosphodiester linkages."
                misc_feature      17..21
                /note="Backbone has phosphorothioate linkages."
                misc_feature      22
                /note="Backbone has phosphodiester linkages."
BASE COUNT      2 a      4 c      14 g      2 t
ORIGIN
Query Match      85.5%; Score 18.8; DB 6; Length 22;
Best Local Similarity 90.9%; Pred. No. 8.7e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 9999gacgatatcgtcggggg 22
    |||||||  |||||||
Db 1 GGGGGACGACGCTGCTCGGGGG 22

RESULT 7
HSM800028      1865 bp      mRNA      linear      PRI 18-FEB-2000
LOCUS      Homo sapiens mRNA; cDNA DKFZ564C103 (from clone DKFZ564C103).
DEFINITION      Homo sapiens mRNA; cDNA DKFZ564C103 (from clone DKFZ564C103).
ACCESSION      AL050269
VERSION      AL050269.1 GI:4886444
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1865)
AUTHORS      Mambut,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE      Direct Submission

```

```

JOURNAL      Submitted (10-MAR-1999) MIPS, Am Klopferpitz 18a, D-82152
            Martinsried, GERMANY
COMMENT
Research Center (UKFZ); Email s.wiemann@kitz-helidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZ564C103) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES
    source      location/Qualifiers
                1..1865
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="DKFZ564C103"
                /tissue_type="brain"
                /clone_lib="564 (synonym: hfpd2). Vector pAMP1; host
                X1-2blue; sites NotI + SalI"
                /dev_stage="fetal"
                49..669
                /gene="DKFZ564C103"
                49..669
                /gene="DKFZ564C103"
                /note="similarity to SPAC577.03c (S.pombe)"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAB43370.1"
                /db_xref="GI:4886445"
                /db_xref="SPTREMBL:O93733"
                /translation="MRLNONTLLGKVVLPYTSHPVRYHEMKSEELORTLASEP
                LTLDEYAMOCWQEDPADKCTFVLDLAEKMOQAPGATFSSCWGDNVLTLDLDTL
                GLEWNIAPSCRGKGLGTENPAMLSYVTTLGLKFRPAKIGQNGEPSTRMQKLHF
                EGVATSSVFQEVTLRLTVSESHQWLEQTSVHEKPYRDSAEPC"
                1838
BASE COUNT      418 a      492 c      572 g      383 t
ORIGIN
polyA_site
Query Match      80.9%; Score 17.8; DB 9; Length 1865;
Best Local Similarity 90.5%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 9999gacgatatcgtcggggg 21
    |||||||  |||||||  |||||||  ||
Db 16 GGGGGACGATTCGCTCGTGG 36

RESULT 8
BC004225      1866 bp      mRNA      linear      PRI 12-JUL-2001
LOCUS      Homo sapiens, DKFZP564C103 protein, clone MGC:4764 IMAGE:3538186,
DEFINITION      mRNA, complete cds.
ACCESSION      BC004225
VERSION      BC004225.1 GI:13278944
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      Strausberg,R.
TITLE      Direct Submission
JOURNAL      Submitted (01-MAR-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.ncl.nih.gov
            Contact: MGC help desk
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DRP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
COMMENT

```

DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
 contact: amadan@systemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 11 Row: 1 Column: 16
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Genomescan gene prediction.

FEATURES

source

Location/Qualifiers
 1. 1886

/organism="Homo sapiens"
 /db_xref="locusid:26151"
 /db_xref="taxon:9606"
 /clone="MGC:4764 IMAGE:3538186"
 /tissue_type="Lung, small cell carcinoma"
 /clone_id="NIH_MGC_7"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 76. 696

CDS

/codon_start=-1
 /product="DKFZP564C103 protein"
 /protein_id="AAH04225.1"
 /db_xref="GI:13278945"
 /translation="MRLNONTLLGKKVLPVPTSEHVPYRHEHMKSEELQRLTASEP
 LTLBOEYAMQCSMOEDADKCTFIVDAEKQAQPGATEESCWGDVNLFTLDELTL
 GEIEVIAEPSCRGKGLGTEAVLAMLSTGTTGLTRFEKKGNGPSIRMFQKLH
 EGVATSSVFEVTLRLTVSESHQWLLEQTSHEKPYRDSAEPC"

BASE COUNT 419 a 496 c 583 g 388 t

ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 1886;
 Best Local Similarity 90.5%; Pred. No. 1.4e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 999ggacatcgtcg999 21
 |||||
 Db 43 GGGGAGCATTCCTCGGTG 63

RESULT 9
 BC004195 1952 bp mRNA linear PRI 12-JUN-2001
 LOCUS Homo sapiens, DKFZP564C103 protein, clone MGC:3586 IMAGE:3528894,
 DEFINITION mRNA, complete cds.
 ACCESSION BC004195
 VERSION BC004195.1 GI:13278863
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1952)
 Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapsb-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
<http://www.systemsbiology.org>
 contact: amadan@systemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia

REMARK

COMMENT

Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 11 Row: c Column: 13
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4886444.
 Location/Qualifiers
 1. 1952

FEATURES

source

/organism="Homo sapiens"
 /db_xref="locusid:26151"
 /db_xref="taxon:9606"
 /clone="MGC:3586 IMAGE:3528894"
 /tissue_type="Muscle, rhadomyosarcoma"
 /clone_id="NIH_MGC_17"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 84. 707

CDS

/codon_start=-1
 /product="DKFZP564C103 protein"
 /protein_id="AAH04195.1"
 /db_xref="GI:13278864"
 /translation="MRLNONTLLGKKVLPVPTSEHVPYRHEHMKSEELQRLTASEP
 LTLBOEYAMQCSMOEDADKCTFIVDAEKQAQPGATEESCWGDVNLFTLDELTL
 GEIEVIAEPSCRGKGLGTEAVLAMLSTGTTGLTRFEKKGNGPSIRMFQKLH
 EGVATSSVFEVTLRLTVSESHQWLLEQTSHEKPYRDSAEPC"

BASE COUNT 476 a 499 c 588 g 389 t

ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 1952;
 Best Local Similarity 90.5%; Pred. No. 1.4e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 999ggacatcgtcg999 21
 |||||
 Db 51 GGGGAGCATTCCTCGGTG 71

RESULT 10
 AF321227 133406 bp DNA linear INV 07-MAR-2001
 LOCUS Tribolium castaneum Ftz (ftz), Scr (scr), Dfd (dfd), Zen (zen), and
 DEFINITION zen2 (zen2) genes, complete cds; and Pz (pz) gene, partial cds.
 ACCESSION AF321227
 VERSION AF321227.1 GI:13241679
 KEYWORDS red flour beetle.
 SOURCE Tribolium castaneum
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 Cucujiformia; Tenebrionidae; Tribolium.
 REFERENCE 1 (bases 1 to 133406)
 Brown, S., Fellers, J., Shippy, T., Denell, R., Stauber, M. and
 Schmidt, O. U.
 TITLE A strategy for mapping bicoid on the phylogenetic tree
 JOURNAL Curr. Biol. 11 (2), R43-R44 (2001)
 MEDLINE 21154823
 REFERENCE 2 (bases 1 to 133406)
 Brown, S.J., Shippy, T.D. and Fellers, J.P.
 TITLE Direct Submission
 JOURNAL Submitted (13-NOV-2000) Division of Biology, Kansas State
 University, Ackert Hall, Manhattan, KS 66506, USA
 LOCATION/Qualifiers
 1. 133406

FEATURES

source

/organism="Tribolium castaneum"
 /strain="GA-2"
 /db_xref="taxon:7070"
 /chromosome="1G 2"
 join(2423..2969,3020..3505)
 /gene="ftz"
 /product="Ftz"
 2423..3505

gene

```
/gene="ftz"
/feature="fushi tarazu; pair-rule gene"
join(2444, .2969,3020, .3366)
/gene="ftz"
/feature="homeodomain transcription factor"
/codon_start=1
/product="ftz"
/protein_id="AAK16421.1"
/db_xref="GI:13241680"
/translation="MNLAKQFSDYRTNSFYQYKTRPFLASDCCKNONTENHTNNE
MTHSVKEEPIYESCRSLNQPYLNHFDNSVTPVNHDFQHLISQYDTCCQGLYPEKT
LINETSLQDENKLDKDDSPALRALTRPQAKTPPMQYENFQOYDNNNSAENS
VSEPPANPEPMKAGDSATGNKRTROTYYTKQLEKEFFNKLTRRRRIELAE
SLRTEROKIWEFONRBMKAKKDKTEQSVSTFPELSLHPAETSMDNVNNDCAKP
LTQINRGPPEPTP"
join(13368, .14001,34099, .34427,34580, .35574)
/mRNA
/gene="scr"
/product="Scr"
<13368, .35574
/gene="scr"
join(13368, .14001,34099, .34403)
/feature="homeodomain transcription factor"
/codon_start=1
/product="Scr"
/protein_id="AAK16422.1"
/db_xref="GI:13241681"
/translation="WSSYTFVNSLAOCGQARDOQSSPECTDIYNNANSTYPPACYS
PPQVAPQYQHPDLYAAPGPHGLQPTMGDTQLOPQRLAPHLQSPWQHPASCTYA
DSTSTGVAHSPQDLSTNGAPRSTPPLSTPSNMTKSSGLTSPLSVSPGKPKATS
TASONLSPASSTSTSTSPERAGTNNNNSSOASSPOIYPMKRVHVGOSTVANG
ETKRORTSTRTOTLEKEFFNKLTRRRRIELAEALCLEROKIKIWFQNRBMKMK
KEHMASNIVYHMSPGYHPQDFLHSPQFHLAT"
join(60024, .61084,68776, .69484)
/mRNA
/gene="dfd"
/product="Dfd"
60024, .69484
/gene="dfd"
join(60418, .61084,68776, .69347)
/feature="homeodomain transcription factor"
/codon_start=1
/product="Dfd"
/protein_id="AAK16423.1"
/db_xref="GI:13241682"
/translation="WSSFLMNPRTALPTYQOPHISGVYVDPKPEPPESEYWNSTYIP
PGDFEGNHIIQOQLOYGYNHONYNSVLSNGYGYVYHPDLHSHGPPIHIOIR
PPLHDTOPMCESLNTQPVYSSNLIQMLADITPVYPCNVKSSVCSGSGHDSR
TSPGSOHLOELGLRLDANSDEDDGDDGOSTNDEDDDEEDGDRQIYPMKRVH
VAGANGTPAPGPEKRORTATYRHQITLEKEFFNKLTRRRRIELAEHLVLSRO
IKIWFONRBMKAKKDKLPNTKVRKTRPVAVTTTKKVRKAKTONATPANNDKPK
SRGEVDGLSENMDMSLTHQISSMGSHAHMHDMLEPMHMGSPSLPLASPGC
GPGGGLOPTISDYGLAL"
join(95847, .95969,96022, .96118,96175, .96788)
/mRNA
/gene="zen"
/product="zen"
95847, .96788
/gene="zen"
join(95874, .95969,96022, .96118,96175, .96722)
/feature="homeodomain transcription factor"
/codon_start=1
/product="zen"
/protein_id="AAK16424.1"
/db_xref="GI:13241683"
/translation="MSYSQFENQAVLOSYNFLOEKTYEYENNOALPPTIYPPSDMN
CPQYNAPYSSVQLQPTVTEIOETOPORTKAGARATATSOVLLEEFHSGK
YLSRPRIOIANLNLSEROIKIWFONRBMKAKKDEMNVSIPRSPAPETASLSPOS
VASTASSDHOIVYRLSHAPITDSANQWTSQITDINSYQCDMLQTSFDCSGTIDMA
LPKQECFTQDECMNQSDFVSPALPILTL"
join(97005, .97187,97243, .97327,97374, .>97984)
/mRNA
/gene="zen2"
/product="zen2"
```

```
gene
<97005, .>97984
/gene="zen2"
join(97005, .97187,97243, .97327,97374, .97984)
CDS
/feature="homeodomain transcription factor"
/codon_start=1
/product="zen2"
/protein_id="AAK16425.1"
/db_xref="GI:13241684"
/translation="MSKDSLLEYKNEEFSDSKLPFSGFLFENQAGNAPVFNNDEN
QALGEKISONSIGNSOMIDYSHPNWSYLTQPEQPFVSLSENSHPPEPRPGA
SNGKRAATVYTSQVLELEERFSKYLCPRPRIQWQMLNTEROKIKIWFONRBMK
KKEEKRVVTPKTSPEASMSPOSTSSNSASBPACOFYLNQFPGSSVYVDETCOY
DEESSYQFNDPQFENYQYNQAVSNVNNYQEGFSQCYGKKNQVGGKVFAGYVSGW
EGQVLENNPQPNITST"
<129720, .>130169
/mRNA
/gene="pb"
/product="pb"
<129720, .>130169
/gene="pb"
129720, .>130169
/feature="homeodomain transcription factor"
/codon_start=1
/product="pb"
/protein_id="AAK16426.1"
/db_xref="GI:13241685"
/translation="MKVHPLNPTGHLDDDFVYKTRNSPEAGGGFWLAAASGA
PHPALPIGCGEFTINSQPSMAERMTALPHLSGMTGASISPDNTPAQYPMQVDP
HMGSGVGVYVPEKPMWKEKTKTSQOQCNFEFRNTVGTATIANIEWRL"
BASE COUNT 44586 a 22191 c 22122 g 44507 t
ORIGIN
Query Match 80.9%; Score 17.8; DB 3; Length 133406;
Best Local Similarity 90.5%; Pred. No. 8.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 gggagacgatacgtcgaggag 22
|||||
Db 102205 GGGGACGATATCGCGGAGG 102225
|||||
RESULT 11
AC016888/c 166044 bp DNA linear HTG 03-SEP-2001
LOCUS
DEFINITION Homo sapiens chromosome 17 clone RP11-45215 map 17, *** SEQUENCING
IN PROGRESS ***, 6 unordered pieces.
ACCESSION AC016888
VERSION AC016888.8 GI:15421997
KEYWORDS HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVERPIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 166044)
Homo sapiens chromosome 17, clone RP11-45215
Unpublished
2 (bases 1 to 166044)
Baldwin,J., Barna,N., Nusbbaum,C., Lander,E., Allen,N., Anderson,M.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Deatellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heathford,A., Horton,L.,
Lehoczky,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,D.,
Hewand,J., Liu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGirk,A., McKernan,K., McLaughlin,J., Meltrin,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Titrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
```

TITLE
JOURNAL
COMMENT

Wyman, D., Ye, W.-J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 3, 2001 this sequence version replaced gi:14140327.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Genome Center
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L5094
Center clone name: 452_I-5

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 38972: contig of 38972 bp in length
38973 39072: gap of 100 bp
39073 45604: contig of 6532 bp in length
45605 45704: gap of 100 bp
45705 82112: contig of 36408 bp in length
82113 82212: gap of 100 bp
82213 98081: contig of 15869 bp in length
98082 98181: gap of 100 bp
98182 105798: contig of 7617 bp in length
105799 105898: gap of 100 bp
105899 166044: contig of 60146 bp in length.

FEATURES

Location/Qualifiers
1. 166044

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"

/clone="RP11-45215"

BASE COUNT 43231 a 41606 c 40776 g 39820 t 611 others

ORIGIN

Query Match

Best Local Similarity 80.9%; Score 17.8; DB 2; Length 166044;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 19999acacatcgtcg999 21

Db 148931 GGGGACGATTCTCGCTG 148911

RESULT 12

AC068874 179594 bp DNA linear HTG 13-FEB-2002
LOCUS Homo sapiens chromosome 17 clone RP11-399J11 map 17, *** SEQUENCING
IN PROGRESS ***, 9 unordered pieces.

ACCESSION

AC068874 AC068874.6 GI:18653759

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 179594)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 17, clone RP11-399J11
unpublished

REFERENCE

AUTHORS

2 (bases 1 to 179594)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barre, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,
Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Deatrelino, K., Dewar, K., Diaz, J.S.,
Dodgson, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Gallagan, J., Gardina, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hayes, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Larocque, K., Lamazares, R., Landers, T., Lenockky, J.,
Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGuck, A., McKernan, R., McPheters, R.,
Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

JOURNAL

COMMENT

Submitted (10-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 13, 2002 this sequence version replaced gi:16931030.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L9436
Center clone name: 399_J-11

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 3340: contig of 3340 bp in length
3341 3440: gap of 100 bp
3441 10827: contig of 7387 bp in length
10828 10927: gap of 100 bp
10928 62109: contig of 51182 bp in length
62110 62209: gap of 100 bp
62210 74525: contig of 12316 bp in length
74526 74625: gap of 100 bp
74626 80741: contig of 6116 bp in length
80742 80841: gap of 100 bp
80842 145765: contig of 64924 bp in length
145766 145865: gap of 100 bp
145866 147976: contig of 2111 bp in length
147977 148076: gap of 100 bp
148077 150266: contig of 2190 bp in length
150267 150366: gap of 100 bp
150367 179594: contig of 29228 bp in length.

Location/Qualifiers

1. 179594

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="17"

/map="17"

/clone="RP11-399J11"

/clone_lib="RP11-399J11 Human Male BAC"

BASE COUNT 40219 a 47157 c 51523 g 39840 t 855 others

ORIGIN

CDS
/gene="lack1 OR SMD20002"
5350..6426
/function="CELL PROCESSES: Transport of small molecules;
carbohydrates, organic acids, alcohols"
/note="Product confidence : probable
Gene name confidence : putative"
/codon_start=1
/transl_table=11
/evidence=not experimental
/product="probable lactose transport ATP-binding protein"
/db_xref="GI:15139877"
/protein_id="CAC48405.1"
/translation="MSELQLSDYKRSYGLGVKVDLIRKSGFVVFVPGGCKST
LRLMAGLEISGDLTIDVPMNDVPSKGIAMVPSQALPHEMVRMNGFALRF
AGVPAIEKRVNAAHILLEGALDKRPGKSGGQROVAIGAIYVRHKIFLDFP
LSNDAELRYHMEIRLKHOKLATITVYTHOVEMTLADKIVYKACVDEGSP
LDLYDDPANLPVAGFISPKMNFLEKGIETDEDDATARLFDYGDAPLPTLQAAAGTA
VTIGIRPHFDEAGPALDLAIDMLEHGEETAYAAHHNGELIVETKNGRLKTG
DRLTARPDVSVLVFDEGKRLS"
6485..7261
/gene="SMD20003"
6485..7261
/function="Small Molecule Metabolism: Amino acid
biosynthesis; proline"
/note="Product confidence : putative
Gene name confidence : hypothetical"
/codon_start=1
/transl_table=11
/evidence=not experimental
/product="putative pyrroline-5-carboxylate reductase
protein"
/protein_id="CAC48406.1"
/db_xref="GI:15139878"
/translation="MRIGFTGTGATITAMTGYGSLDVAGILVSPRNRIANRLAG
RPLPLRVKNDQAVADVLAVRQALIEDVLPKLFKRGQIVSVIATDRSTLL
GMIGEEVELTQALPLPFAVEGVTATYPPNRAVVAALFAALGTAVECTREEDLLAV
GSAIMATYLGILDRITWFAEKGLHRDKARAYATAPLASLAGRAVRDEGTPLEELARE
FSTRKGLENEVLDFDRGGHGLTDLAENVLLRVRRDR"
complement(7268..8173)
/gene="gscr OR SMD20004"
complement(7268..8173)
/gene="gscr OR SMD20004"
/function="MISCELLANEOUS"
/note="Product confidence : putative
Gene name confidence : putative"
/codon_start=1
/transl_table=11
/evidence=not experimental
/product="putative gstr-like transcriptional regulator
protein"
/protein_id="CAC48407.1"
/db_xref="GI:15139879"
/translation="MDRMQAMRIFFOVYSGSGFAPAKVILHMSPPSVTRAVALLEDLI
GRLVTRTSRLKLTAAEGYVADCRILGEIEBAENANAGSTAPAGLITVAPALF
GRILHLPVILDFDHYPAWOKTIFVDRVTNLVDEGLVARIASLPSGLVARIS
VRQVLCGSPDFARFGDPSGPELARIRIIRGGLSHSMLFGRONINIVPSPLRI
CMTNDAAVIAAVAGLSPROSYOVAPDSAGLKYVLADYEPEPVIHIVHAGRIY
SARVRAVDVRAAFRRRHAGLRTLE"
8289..8900
/gene="gstl5 OR SMD20005"
8289..8900
/gene="gstl5 OR SMD20005"
/EC_number="2.5.1.18"
/function="MISCELLANEOUS"
/note="Product confidence : putative
Query Match 80.9%; Score 17.8; DB 1; Length 310050;
Best Local Similarity 90.5%; Pred. No. 7.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 gggagcagatcgtcg9999 22

Db 43529 GGGACATATCATCGGGG 43549
||||| ||||| |||||
RESULT 14
AX346175
LOCUS AX346175 8718 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 1246 from Patent WO0200928.
ACCESSION AX346175
VERSION AX346175.1 GI:18494061
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..8718
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 2225 a 183 c 2124 g 4186 t
ORIGIN
Query Match 79.1%; Score 17.4; DB 6; Length 8718;
Best Local Similarity 94.7%; Pred. No. 1.7e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 gggagcagatcgtcg999 20
||||| ||||| |||||
Db 3302 GGGGCGCATATCGTCGGG 3320
RESULT 15
OSU31773/c
LOCUS OSU31773 954 bp mRNA linear PLN 23-MAR-1999
DEFINITION Oryza sativa protein phosphatase 1 mRNA, complete cds.
ACCESSION U31773
VERSION U31773.1 GI:951335
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
rice.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 954)
Chang, M., Wang, B., Chen, X. and Wu, R.
Molecular characterization of catalytic-subunit cDNA sequences
encoding protein phosphatases 1 and 2A and study of their roles in
the gibberellin-dependent Osamy-c expression in rice
Plant Mol. Biol. 39 (1), 105-115 (1999)
99178798
2 (bases 1 to 954)
Wang, B., Chang, M., Chen, X. and Wu, R.
Direct Submission
Submitted (18-JUL-1995) Baiyang Wang, Biochemistry, Cornell
University, Ithaca, NY 14853, USA
location/Qualifiers
1..954
/organism="Oryza sativa"
/db_xref="taxon:4530"
/tissue_type="seed"
1..954
/codon_start=1
/product="protein phosphatase 1"
/protein_id="AAA74625.1"
/db_xref="GI:951336"
/translation="MAPRRGGGMPVLLDDIIPPLLEVRTARPGQVHLSESEINQL
CSVSRIFLSQPNLLELPAIKICGDIHQYSDLLRFLFYGGPPPEANVLFGLDYDR

BASE COUNT 242 a 214 c 233 g 265 t

Query Match	78.2%	Score 17.2	DB 8	Length 954
Best Local Similarity	86.4%	Pred. No. 2.7e+03		
Matches 19	Conservative 0	Mismatches 3	Indels 0	Gaps 0
Oy	1	ggggagacatactctcggggg	22	
Db	61	ggggagatcattacgtcgcaggag	40	

Search completed: August 10, 2002, 02:57:51
Job time: 15677 sec

THIS PAGE BLANK (COPY)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2002, 22:38:54 ; Search time 277.54 Seconds
(Without alignments)
17.701 Million cell updates/sec

Title: US-09-672-126-7

Perfect score: 20

Sequence: 1 gggggagcagtcgctgggggg 20

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16.4	82.0	1776	1	US-08-522-229B-1
2	16.4	82.0	1776	2	US-09-102-433-1
3	15.2	76.0	882	4	US-08-818-112-138
4	15.2	76.0	882	4	US-08-818-111-133
5	15.2	76.0	882	4	US-09-056-556-138
6	15.2	76.0	1416	1	US-08-236-311-3
7	15.2	76.0	1416	3	US-08-457-918-3
8	15.2	76.0	1508	1	US-08-236-311-6
9	15.2	76.0	1508	3	US-08-457-918-6
10	15.2	76.0	4403765	4	US-09-103-840A-2
11	14.4	72.0	1207	1	US-08-460-806-16
12	14.4	72.0	1207	1	US-08-325-630-16
13	14.2	71.0	25	4	US-08-853-774-19
14	14.2	71.0	826	4	US-08-853-774-4
15	14.2	71.0	1694	4	US-09-362-473-3
16	14.2	71.0	3429	1	US-08-097-997A-10
17	14.2	71.0	3429	3	US-08-665-574C-10
18	14.2	71.0	3429	4	US-08-946-994-10
19	14.2	71.0	4234	1	US-08-446-038B-1
20	14.2	71.0	4234	1	US-08-446-010B-1
21	14.2	71.0	4234	1	US-08-805-445-1
22	14.2	71.0	4234	2	US-08-064-067D-1
23	14.2	71.0	4234	2	US-09-066-208-1
24	14.2	71.0	8438	1	US-07-945-283-1
25	14.2	71.0	11604	4	US-09-385-028-13
26	14.2	71.0	15079	4	US-09-385-028-1
27	14.2	71.0	30001	1	US-08-125-468-1

C	28	14.2	71.0	30001	1	US-08-125-468-1	Sequence 1, Appl
C	29	14.2	71.0	30001	2	US-08-474-933-1	Sequence 1, Appl
C	30	14.2	71.0	30001	2	US-08-474-933-1	Sequence 1, Appl
C	31	13.8	69.0	1389	2	US-08-023-880B-3	Sequence 3, Appl
C	32	13.8	69.0	1389	2	US-08-486-953A-3	Sequence 3, Appl
C	33	13.8	69.0	1396	6	5472691-1	Patent No. 5472691
C	34	13.8	69.0	1406	1	US-08-745-769-1	Sequence 1, Appl
C	35	13.8	69.0	1406	2	US-08-157-185-1	Sequence 1, Appl
C	36	13.8	69.0	1406	3	US-08-281-526B-1	Sequence 1, Appl
C	37	13.8	69.0	1406	4	US-09-450-797-1	Sequence 1, Appl
C	38	13.8	69.0	1406	5	PCT-US93-10553-1	Sequence 1, Appl
C	39	13.8	69.0	1417	2	US-08-428-243-8	Sequence 8, Appl
C	40	13.8	69.0	1417	5	PCT-US93-10301-8	Sequence 8, Appl
C	41	13.8	69.0	1702	1	US-07-616-022C-1	Sequence 1, Appl
C	42	13.8	69.0	2459	1	US-08-101-593-5	Sequence 5, Appl
C	43	13.8	69.0	2459	1	US-08-465-995A-5	Sequence 5, Appl
C	44	13.8	69.0	2459	2	US-08-465-994C-5	Sequence 5, Appl
C	45	13.8	69.0	3661	4	US-08-718-388-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-522-229B-1
; Sequence 1, Application US/08522229B
; Patent No. 5811291
; GENERAL INFORMATION:
; APPLICANT: Kolod, Lene Venke
; APPLICANT: Andersen, Lene No. 5811291boe
; APPLICANT: Dalboge, Henrik
; APPLICANT: Kauppinen, Markus Sakari
; APPLICANT: Christgau, Stephen Peter
; APPLICANT: Heidt-Hansen, Hans Peter
; APPLICANT: Christopher, Claus
; APPLICANT: Nielsen, Per Munk
; APPLICANT: Voragen, Alphons Gerard Joseph
; TITLE OF INVENTION: An Enzyme With Rhamnogalacturonase Activity
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 58112910 No. 5811291disk of No. 5811291th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/522,229B
; FILING DATE: 29-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valetta
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 3953, 204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 867-0123
; TELEFAX: (212) 878-9655
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1776 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 7...1587
; OTHER INFORMATION:

NAME/KEY: Signal Sequence
LOCATION: 7...63
OTHER INFORMATION:
US-08-522-229B-1

Query Match 82.0%; Score 16.4; DB 1; Length 1776;
Best Local Similarity 94.4%; Pred. No. 18;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 999gacgatcgtcg999g 19
|||||
DB 1476 GGGGACGATCGTCGGCGG 1493

RESULT 2

US-09-102-433-1
Sequence 1, Application US/09102433
Patent No. 5882911

GENERAL INFORMATION:

APPLICANT: Kofoed, Lene Venke
APPLICANT: Andersen, Lene No. 5882911boe
APPLICANT: Dalboge, Henrik
APPLICANT: Kauppinen, Markus Sakari
APPLICANT: Christgau, Stephen
APPLICANT: Heidt-Hansen, Hans Peter
APPLICANT: Christoffersen, Claus
APPLICANT: Nielsen, Per Munk
APPLICANT: Voragen, Alphons Gerard Joseph
TITLE OF INVENTION: An Enzyme With Rhamnogalacturonase Activity
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 5882911disk of No. 5882911th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,433
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/522,229
FILING DATE: 29-AUG-1995

ATTORNEY/AGENT INFORMATION:

NAME: Gregg, Valeta
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 3953,204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 867-0123
TELEFAX: (212) 878-9655

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1776 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA

FEATURE:

NAME/KEY: Coding Sequence
LOCATION: 7...1587

OTHER INFORMATION:

NAME/KEY: Signal Sequence
LOCATION: 7...63
OTHER INFORMATION:

US-09-102-433-1

Query Match 82.0%; Score 16.4; DB 2; Length 1776;
Best Local Similarity 94.4%; Pred. No. 18;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 999gacgatcgtcg999g 19
|||||
DB 1476 GGGGACGATCGTCGGCGG 1493

RESULT 3

US-08-818-112-138/c
Sequence 138, Application US/08818112
Patent No. 629096

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,112
FILING DATE: 13-MAR-1997
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121,411c6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 138:

SEQUENCE CHARACTERISTICS:
LENGTH: 882 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-818-112-138

Query Match 76.0%; Score 15.2; DB 4; Length 882;
Best Local Similarity 85.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 999gacgatcgtcg999g 20
|||||
DB 579 GGGGACGATCGTCGGCGG 560

RESULT 4

US-08-818-111-133/c
Sequence 133, Application US/08818111
Patent No. 633852

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.

	Query Match	76.0%;	Score 15.2;	DB 4;	Length 882;
	Best Local Similarity	85.0%;	Pred. No. 63;		
	Matches 17;	Conservative	0;	Mismatches 3;	Indels 0;
Qy	1	gggggacagatcgtcgggggg	20		
Db	579	ggggcaccgcgcgcgcggggcg	560		

; ZIP: 94080

APPLICATION NUMBER: US/08/236,311

FILING DATE: 02-MAY-1994

APPLICATION NUMBER: 07/036100

PRIOR APPLICATION DATA: 07/04/07

ENDING DATE: 10 FEB 1992

PRIOR APPLICATION DATA:

AFILIATION NUMBER: 07/200703
 EXPIRATION DATE: 30-SEP-1088

APPLICATION NUMBER: 07/106329

FILED DATE: 02-OCT-1987

NAME: Hasak, Janet E.

REGISTRATION NUMBER: 28,616

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/223 103
TELEFAX: 415/953-9891

; CURRENT APPLICATION DATA:

TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1416 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-236-311-3

Query Match 76.0%; Score 15.2; DB 1; Length 1416;
Best Local Similarity 85.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggggagcgtcgtcgggggg 20
||||| ||| ||||| |
Db 515 GGGGACCATCATCGGGAG 534

RESULT 7
US-08-457-918-3
Sequence 3, Application US/08457918
Patent No. 6117655
GENERAL INFORMATION:
APPLICANT: Capon, Daniel J.
APPLICANT: Gregory, Timothy J.
TITLE OF INVENTION: Adhesion Variants
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,918
FILING DATE: 1-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/236311
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/936190
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842777
FILING DATE: 18-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/250785
FILING DATE: 28-SEP-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/104329
FILING DATE: 02-OCT-1987
ATTORNEY/AGENT INFORMATION:
NAME: Kubienc, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0444PIC3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8228
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1416 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-457-918-3

Query Match 76.0%; Score 15.2; DB 3; Length 1416;
Best Local Similarity 85.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggggagcgtcgtcgggggg 20
||||| ||| ||||| |
Db 515 GGGGACCATCATCGGGAG 534

RESULT 8
US-08-236-311-6
Sequence 6, Application US/08236311
Patent No. 5565335
GENERAL INFORMATION:
APPLICANT: Capon, Daniel J.
APPLICANT: Gregory, Timothy J.
TITLE OF INVENTION: Adhesion Variants
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/236,311
FILING DATE: 02-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/936190
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842777
FILING DATE: 18-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/250785
FILING DATE: 28-SEP-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/104329
FILING DATE: 02-OCT-1987
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 444PIC2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1508 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-236-311-6

Query Match 76.0%; Score 15.2; DB 1; Length 1508;
Best Local Similarity 85.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggggagcgtcgtcgggggg 20
||||| ||| ||||| |
Db 607 GGGGACCATCATCGGGAG 626

RESULT 9
US-08-457-918-6
Sequence 6, Application US/08457918
Patent No. 6117655
GENERAL INFORMATION:
APPLICANT: Capon, Daniel J.
APPLICANT: Gregory, Timothy J.
TITLE OF INVENTION: Adheson Variants
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 KB floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: palin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,918
FILING DATE: 1-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/236311
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/936190
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842777
FILING DATE: 18-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/250785
FILING DATE: 28-SEP-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/104329
FILING DATE: 02-OCT-1987
ATTORNEY/AGENT INFORMATION:
NAME: Kudinec, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0444PIC3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/952-8228
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1508 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-457-918-6

Query Match 76.0%; Score 15.2; DB 3; Length 1508;
Best Local Similarity 85.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gggggagcagtcgctggggggg 20
||||| ||| |||||
Db 607 GGGGACCATCATCGGGGAG 626

RESULT 10
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:

APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 76.0%; Score 15.2; DB 4; Length 4403765;
Best Local Similarity 85.0%; Pred. No. 41;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gggggagcagtcgctggggggg 20
||||| ||| |||||
Db 3700683 GGTGACGATCGTCGGGCG 3700664

RESULT 11
US-08-460-806-16/c
Sequence 16, Application US/08460806
Patent No. 5747241
GENERAL INFORMATION:
APPLICANT: MIYAMURA, TATSUO
APPLICANT: SAITO, IZUMU
APPLICANT: HARADA, SHIZUKO
APPLICANT: HONDA, YOSHIKAZU
TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS C
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,806
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/325,630
FILING DATE: 19-OCT-1994
APPLICATION NUMBER: US 07/956,993
FILING DATE: 06-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5747241man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 4667-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 16:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C virus
; IMMEDIATE SOURCE:
; CLONE: H90
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1207
; US-08-460-806-16

Query Match          72.0%; Score 14.4; DB 1; Length 1207;
Best Local Similarity 93.8%; Pred. No. 1.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 gggacgctcgcggg 18
Db 436 GGGACGCTCTCGGG 421

RESULT 12
US-08-325-630-16/c
; Sequence 16, Application US/08325630
; Patent No. 5750331
; GENERAL INFORMATION:
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: SAITO, IZUMU
; APPLICANT: HARADA, SHIZUKO
; APPLICANT: HONDA, YOSHIKAZU
; TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS C
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,630
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/956,993
; FILING DATE: 06-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NO. 5750331man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 4667-001-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
```

```

; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C virus
; IMMEDIATE SOURCE:
; CLONE: H90
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1207
; US-08-325-630-16

Query Match          72.0%; Score 14.4; DB 1; Length 1207;
Best Local Similarity 93.8%; Pred. No. 1.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 gggacgctcgcggg 18
Db 436 GGGACGCTCTCGGG 421

RESULT 13
US-08-853-774-19/c
; Sequence 19, Application US/08853774
; Patent No. 6265557
; GENERAL INFORMATION:
; APPLICANT: Diamond, David
; APPLICANT: Nehlsen-Cannarella, Sandra
; APPLICANT: Fagoaga, Omar
; APPLICANT: Szalay, Aladar
; TITLE OF INVENTION: ABO HISTO-BLOOD GROUP O ALLELES OF THE BABOON
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive Sixteenth Flo
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,774
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: LOMAIMM.100A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714/760-0404
; TELEFAX: 714/760-9503
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-853-774-19

Query Match          71.0%; Score 14.2; DB 4; Length 25;
Best Local Similarity 84.2%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gggacgctcgcggg 20
Db 23 GGGGCGCTCTCGGG 5
```


RESULT 14

US-08-853-774-4

; Sequence 4, Application US/08853774

; Patent No. 6265557

; GENERAL INFORMATION:

; APPLICANT: Diamond, David

; APPLICANT: Nehlsen-Cannarella, Sandra

; APPLICANT: Paooga, Omar

; APPLICANT: Salay, Aladar

; TITLE OF INVENTION: ABO HISTO-BLOOD GROUP O ALLELES OF THE BABOON

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobb, Martens, Olson & Bear

; STREET: 620 Newport Center Drive Sixteenth Flo

; CITY: Newport Beach

; STATE: CA

; COUNTRY: USA

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/853,774

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Altman, Daniel E

; REGISTRATION NUMBER: 34,115

; REFERENCE/DOCKET NUMBER: LOMAIMM.100A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 714/760-0404

; TELEFAX: 714/760-9503

; TELEX:

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 826 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-853-774-4

Query Match

Best Local Similarity 71.0%; Score 14.2; DB 4; Length 826;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 999gacgacgtcggggg 20

||||| ||| ||| ||| |||

Db 559 GGGGGCGTCTCTCGGGGG 577

RESULT 15

US-09-362-473-3/C

; Sequence 3, Application US/09362473

; Patent No. 6218169

; GENERAL INFORMATION:

; APPLICANT: Cahoon, Edgar B.

; APPLICANT: Cahoon, Rebecca E.

; APPLICANT: Falco, S. Carl

; APPLICANT: Morgante, Michele

; APPLICANT: Rafalski, J. Antoni

; APPLICANT: Hitz, William D.

; APPLICANT: Kinney, Anthony J.

; TITLE OF INVENTION: Aromatic Amino Acid Catabolism Enzymes

; FILE REFERENCE: BB-1197

; CURRENT APPLICATION NUMBER: US/09/362,473

; CURRENT FILING DATE: 1999-07-28

; EARLIER APPLICATION NUMBER: 60/094,783

; EARLIER FILING DATE: JULY 31, 1998

; NUMBER OF SEQ. ID NOS: 14

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 3

; LENGTH: 1694

; TYPE: DNA

; ORGANISM: Oryza sativa

; US-09-362-473-3

Query Match

Best Local Similarity 71.0%; Score 14.2; DB 4; Length 1694;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 999gacgacgtcggggg 20

||||| ||| ||| ||| |||

Db 477 GGGGAGGATGTTGGGGGG 459

Search completed: August 10, 2002, 03:03:41
Job time: 15887 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2002, 22:32:19 ; Search time 9068.22 Seconds
(without alignments)
29.768 Million cell updates/sec

Title: US-09-672-126-7

Perfect score: 20

Sequence: 1 gggggagcagcagcgcggggg 20

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.4	92.0	492	10	BG789447 6HRm46 6H
2	17.4	87.0	961	10	BE959048 601644811
3	16.8	84.0	247	10	BE951215 EST18475
4	16.8	84.0	567	10	B0134969 B0134969
5	16.8	84.0	580	12	CNS04BH3
6	16.8	84.0	612	12	AQ272353 Tetradon
7	16.8	84.0	643	10	AQ272353 nbx00027J
8	16.8	84.0	762	10	B0140169 B0140169
9	16.8	84.0	898	10	BG665983 602788065
10	16.8	84.0	994	10	BE960471 601653215
11	16.8	84.0	1030	12	AG136593
12	16.8	84.0	1081	10	BM470460
13	16.8	84.0	1101	12	CNS05SUN
14	16.8	84.0	1120	10	BE965826 601659002
15	16.8	84.0	1347	10	BM480133 AGENCOURT
16	16.8	84.0	1427	10	BG167937 602340029
17	16.8	84.0	1758	10	BE963644 601656656

18	16.4	82.0	461	10	W79641
19	16.4	82.0	551	9	AW745219
20	16.4	82.0	688	9	AA753250
21	16.4	82.0	923	10	BF036664 601459743
22	16.4	82.0	1005	12	AG131408
23	15.8	79.0	169	10	H54705
24	15.8	79.0	284	9	BB089505
25	15.8	79.0	291	9	AV096501
26	15.8	79.0	435	10	B1337090
27	15.8	79.0	482	10	BG463924
28	15.8	79.0	483	9	AA578972
29	15.8	79.0	492	10	BG948810
30	15.8	79.0	496	9	A1622333
31	15.8	79.0	516	10	BM429036
32	15.8	79.0	519	10	BE479770
33	15.8	79.0	573	9	A1670200
34	15.8	79.0	581	10	BF041244
35	15.8	79.0	627	10	BF046655
36	15.8	79.0	633	10	BE584875
37	15.8	79.0	637	10	B1954807
38	15.8	79.0	647	9	A1670160
39	15.8	79.0	647	9	AL661185
40	15.8	79.0	689	10	BG701743
41	15.8	79.0	738	12	A2209087
42	15.8	79.0	744	10	BG837757
43	15.8	79.0	749	12	AG072754
44	15.8	79.0	761	10	BG837526
45	15.8	79.0	763	10	BE368455

ALIGNMENTS

RESULT 1
LOCUS BG789447/c
DEFINITION 6HRm46 6HR Nitrogen-limited Schizophyllum library
Schizophyllum commune cdna 5' similar to mannose-1-phosphate guanylyltransferase, mRNA sequence.
ACCESSION BG789447
VERSION BG789447.1 GI:14124998
KEYWORDS EST
SOURCE Schizophyllum commune
ORGANISM Schizophyllum commune
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes; Agaricales; Schizophyllaceae; Schizophyllum
REFERENCE 1 (bases 1 to 492)
Guettler, S., Lucchese, S.A., Honaas, L.A., Hittinger, C.T., Green, A., Lilly, M.W. and Gathman, A.C.
More expressed sequence tags from Schizophyllum commune nitrogen-replete and nitrogen-limited libraries
Unpublished (2001)
JOURNAL Contact: Gathman AC
Biology Department
Southeast MO State University
1 University Plaza, Cape Girardeau, MO 63701, USA
Tel: 5736512361
Fax: 5739866433
Email: agathman@biology.smo.edu
Seq primer: T3
POLYA-No.
FEATURES
source location/Qualifiers
1..492
/organism="Schizophyllum commune"
/strain="4-40"
/db_xref="taxon:5334"
/clone_lib="6HR Nitrogen-limited Schizophyllum library"
/tissue_type="mycelium"
/note="Vector: lambda Zap; Site_1: EcoRI; Site_2: XhoI; 4 day-old mycelia of Schizophyllum commune were transferred from minimal (nitrogen-replete) medium to low-nitrogen medium. RNA was extracted six hours after transfer and cDNAs prepared."

BASE COUNT 71 a 189 c 128 g 102 t 2 others
ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 492;
Best Local Similarity 95.0%; Pred. No. 5.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 9999gacatcgtcg9999 20
|||||
Db 35 GGGGAGCAGATGTCGGGG 16

RESULT 2
LOCUS BE959048 961 bp mRNA linear EST 04-OCT-2000
DEFINITION 60164481R2 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:3929859 3',
mRNA sequence.
ACCESSION BE959048 GI:10569753
VERSION BE959048.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
1 (bases 1 to 961)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LCM763 row: f column: 04
High quality sequence stop: 1.

FEATURES
source location/Qualifiers

1..961
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3929859"
/clone_1lb="NIH_MGC_56"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Brain; Vector: pDNR-LIB (Clontech); Site_1:
Site_2: Site_3: Site_4: Site_5: Site_6: Site_7: Site_8:
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCTAGAGCGGAGCGGCGGCACATG-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 229 a 217 c 276 g 239 t
ORIGIN

Query Match 87.0%; Score 17.4; DB 10; Length 961;
Best Local Similarity 94.7%; Pred. No. 1.7e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 9999gacatcgtcg9999 19
|||||
Db 37 GGGGAGCAGATGTCGGGG 55

RESULT 3

BF251215/c 247 bp mRNA linear EST 15-NOV-2001
LOCUS BF251215
DEFINITION EST418475 Coccidioides immitis spherule cDNA library Coccidioides
immitis cDNA clone CIAAE36 5' sequence, mRNA sequence.
ACCESSION BF251215
VERSION BF251215.1 GI:16931358
KEYWORDS EST.
SOURCE Coccidioides immitis.
ORGANISM Coccidioides immitis

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Orygenales; mitosporic Orygenales; Coccidioides.
AUTHORS Gardner, M.J. and Kirkland, T.
1 (bases 1 to 247)
TITLE Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org.

FEATURES
source location/Qualifiers

1..247
/organism="Coccidioides immitis"
/db_xref="taxon:5501"
/clone_image="CIAAE36"
/clone_1lb="Coccidioides immitis spherule cDNA library"
/dev_stage="spherule"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 61 a 78 c 44 g 64 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 247;
Best Local Similarity 90.0%; Pred. No. 2.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 9999gacatcgtcg9999 20
|||||
Db 96 GGGGAGCAGATGTCGGGG 77

RESULT 4
LOCUS B1134969 567 bp mRNA linear EST 23-JAN-2002
DEFINITION B1134969 unpublished oligo-capped cDNA library, C. elegans L1 stage
Caenorhabditis elegans cDNA clone yk1094e12 3', mRNA sequence.
ACCESSION B1134969
VERSION B1134969.1 GI:18295126
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditiodea
1 (bases 1 to 567)
AUTHORS Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.

and Sugano, S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished (2002)
COMMENT Contact: Tadashi Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6855
Fax: 81-559-81-6856
Email: tshin@genes.nig.ac.jp.

FEATURES
source location/Qualifiers
1..567
/organism="Caenorhabditis elegans"
/strain="N2"

	Query Match	84.0%; Score 16.8; DB 10; Length 567;
	Best Local Similarity	90.0%; Pred. No. 2.6e+03;
Matches	18; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
OY	1 ggggagacatcgtcgaggag 20 Db 349 GGGGAGCAGTGTGGGGCGG 368	
RESULT 5	CNS04BH3	580 bp DNA linear GSS 21-MAY-2000
LOCUS	Tetraodon nigroviridis genome survey sequence T7 end of clone 097020 of library G from Tetraodon nigroviridis, genomic survey sequence.	
DEFINITION	AT283152.1 GI:8021509	
ACCESSION	AL283152.1	GI:8021509
VERSION	GSS: genome survey sequence.	
KEYWORDS	Tetraodon nigroviridis	
SOURCE	Tetraodon nigroviridis	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.	
REFERENCE	1 (bases 1 to 580)	
AUTHORS	Roeest-Crollius,H., Jalllon,O., Dasllva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.	
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis	
JOURNAL REFERENCE	Unpublished	
AUTHORS	2 (bases 1 to 580)	
TITLE	Roeest-Crollius,H., Jalllon,O., Dasllva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottler,P., Quetier,F., Saurin,W. and Weissenbach,J.	
JOURNAL REFERENCE	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence	
AUTHORS	Unpublished	
TITLE	3 (bases 1 to 580)	
JOURNAL REFERENCE	Genoscope.	
AUTHORS	Direct Submission	
TITLE	Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases	
JOURNAL COMMENT	This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.	
FEATURES	Location/Qualifiers	
source	1..580	/organism="Tetraodon nigroviridis"
	/db_xref="taxon:99883"	
	/clone="097020"	
	/clone_id="G"	
	/note="Genoscope sequence ID : COBG097BH10LP1-end : T7"	
BASE COUNT	116 a 124 c 130 g 176 t 34 others	
ORIGIN		
Query Match	84.0%; Score 16.8; DB 12; Length 580;	
Best Local Similarity	90.0%; Pred. No. 2.7e+03;	
Matches	18; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
OY	1 ggggagacatcgtcgaggag 20	

```

RESULT      6
LOCUS       AO272353/c
DEFINITION  nbxb00027101or CUG1 Rice BAC Library Oryza sativa genomic clone
ACCESSION   AO272353
VERSION     AO272353.1
KEYWORDS    GI:3825668
SOURCE      Oryza sativa.
ORGANISM    Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 612)
Wing R.A. and Dean R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing R
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAAACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 360.
Location/Qualifiers
1. 612
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbxb00270101or"
/clone_1ib="CUG1 Rice BAC Library"
/lssue_type="Leaf"
/lab_host="E. coli DH10B"
/note="Vector: pBeloBAC11; Site 1: HindIII; Site 2:
HindIII. Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 Kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9 %. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."
BASE COUNT  161 a 135 c 120 g 195 t 1 others
ORIGIN
Query Match      84.0%; Score 16.8; DB 12; Length 612;
Best Local Similarity 90.0%; Pred.No. 2.7e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Ory 1 gggggagcagctcgcgggg 20
Db 547 gggggagcagctcgcgggg 528

```

```

RESULT 7
BJ140169      643 bp      mRNA      linear      EST 23-JAN-2002
LOCUS        BJ140169 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION   Caenorhabditis elegans cDNA clone YK155a05 3', mRNA sequence.
ACCESSION    BJ140169
VERSION      BJ140169.1 GI:18300335
KEYWORDS     EST.
SOURCE       Caenorhabditis elegans.
ORGANISM     Caenorhabditis elegans.
REFERENCE    Bkaryota: Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea
AUTHORS      (1 bases 1 to 643)
              Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
              and Sugano, S.
TITLE        A complementary view of the C. elegans genome
JOURNAL      Unpublished (2002)
COMMENT      Contact: Tadasi Shin-I
              Center For Genetic Resource Information
              National Institute of Genetics
              1111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tsun1@genes.nig.ac.jp.
              Location/Qualifiers
FEATURES
    source
        1..643
        /organism="Caenorhabditis elegans"
        /strain="N2"
        /db_xref="taxon:62339"
        /clone_lib="unpublished oligo-capped cDNA library, C.
        elegans L1 stage"
        /sex="hermaphrodite"
        /tissue_type="whole animal"
        /dev_stage="L1"
BASE COUNT   222 a      85 c      153 g      179 t      4 others
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 643;
Best Local Similarity 90.0%; Pred. No. 2.7e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 99999acgacgtcgtcggggg 20
    |||||  |||||  |||||  |||||  |||||
Db 358 GGGGGGAGATGGTGGGGGG 377

RESULT 8
BG865983      762 bp      mRNA      linear      EST 29-MAY-2001
LOCUS        BG865983
DEFINITION   602788065F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4913945 5',
ACCESSION    BG865983
VERSION      BG865983.1 GI:14216523
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus.
REFERENCE    Mammalia: Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
AUTHORS      1 (bases 1 to 762)
              NIH-MGC http://mgc.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-remail.nih.gov
              Tissue Procurement: Jeffrey E. Green, M.D.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov

```

```

Plate: LLAM10820 row: a column: 18
High quality sequence start: 4
High quality sequence stop: 371.
Location/Qualifiers
1..762
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="IMAGE:4913945"
/lab_host="NCI_CGAP_SG2"
/clone_lib="NCI_CGAP_SG2"
/clone_lib="DHI0B (71 phage-resistant)"
/Note="Organ: salivary gland; Vector: pCMV-Sport6; Site: 1;
NotCl; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo
dr. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT   137 a      368 c      159 g      98 t
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 762;
Best Local Similarity 90.0%; Pred. No. 2.8e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 99999acgacgtcgtcggggg 20
    |||||  |||||  |||||  |||||  |||||
Db 707 GGGGGGCGGTGTCGGGGGG 688

RESULT 9
BE960471      898 bp      mRNA      linear      EST 04-OCT-2000
LOCUS        BE960471
DEFINITION   601653215R2 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:3826371 3',
ACCESSION    BE960471
VERSION      BE960471.1 GI:10571176
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Homo.
              1 (bases 1 to 898)
              NIH-MGC http://mgc.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: CLONTECH Laboratories, Inc.
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM493 row: n column: 04.
              Location/Qualifiers
FEATURES
    source
        1..898
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone_lib="IMAGE:3826371"
        /clone_lib="NIH_MGC_58"
        /tissue_type="hypertrophoma"
        /lab_host="DHI0B (71 phage-resistant)"
        /Note="Organ: kidney; Vector: pNR-LIB (Clontech); Site: 1;
        SfiI (ggcgccgtggcc); Site: 2; SfiI (ggcattatggcc);
        Double-stranded cDNA was prepared from cell line RNA. 5'
        and 3' adaptors were used in cloning as follows: 5'
        adaptor sequence: 5'-CACGGCCATTATGACC-3' and 3' adaptor
        sequence: 5'-ATTCAGAGCGCGAGCGCGCATG-dt(30)BN-3'
        (where B = A, C, G or T and N = A, C, G or T). Average
        insert size 1.35 kb (range 0.9-4.0 kb). 15/15 clones
        contained inserts by PCR. This library was enriched for
        full-length clones and was constructed by Clontech
        Laboratories (Palo Alto, CA)."
```

Sat Aug 10 09:08:42 2002

us-09-672-126-7.1st

```

BASE COUNT      190 a      227 c      333 g      147 t      1 others
ORIGIN

Query Match      84.0%; Score 16.8; DB 10; Length 898;
Best Local Similarity 90.0%; Pred. No. 2.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 9999gacatcgtcg9999g 20
      11111 11111 11111
Db      12 GGGGGGCGATCGCGGGG 31

RESULT 10      994 bp      mRNA      linear      EST 13-FEB-2001
BG243252      BG243252      602355673F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4483866.5',
LOCUS      mRNA sequence.
DEFINITION      BG243252
ACCESSION      BG243252
VERSION      BG243252.1 GI:12753067
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
      Eukaryota; Metazoa; Rodentia; Scurionath; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 994)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
      National Institutes of Health, Mammalian Gene Collection (MGC)
      Unpublished (1999)
      Contact: Robert Strausberg, Ph.D.
      Email: cgabbs-remail.nih.gov
      Tissue Procurement: Gilbert Smith, Ph.D.
      CDNA Library Preparation: Life Technologies, Inc.
      DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
      DNA Sequencing by: Incyte Genomics, Inc.
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LNL at:
      http://image.llnl.gov
      Plate: L1AM10323 row: a column: 19
      High quality sequence stop: 675.
      Location/Qualifiers
FEATURES
      source
      1..994
      /organism="Mus musculus"
      /strain="FVB/N"
      /db_xref="taxon:10090"
      /clone="IMAGE:4483866"
      /clone_lib="NCI_CGAP_Mam1"
      /tissue_type="tumor, biopsy sample"
      /dev_stage="10 months, virgin"
      /lab_host="DH10B"
      /note="Organ: mammary; Vector: pCMV-SPORT6; site_1: SalI;
      site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
      Library constructed by Life Technologies. Investigator
      providing samples: Gilbert Smith, NIH"
BASE COUNT      241 a      254 c      302 g      197 t

Query Match      84.0%; Score 16.8; DB 10; Length 994;
Best Local Similarity 90.0%; Pred. No. 3e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 9999gacatcgtcg9999g 20
      11111 11111 11111
Db      782 GGGGGGCGATCGCGGGG 801

RESULT 11      1030 bp      DNA      linear      GSS 04-NOV-2001
AG136593      AG136593      pan troglodytes DNA, clone: PTB-150D13.F, genomic survey sequence.
LOCUS      AG136593
DEFINITION      AG136593
ACCESSION      AG136593.1 GI:1666271
VERSION

```

```

KEYWORDS      GSS (genome survey sequence).
SOURCE      pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
      BAC Library clone:PTB-150D13.F.
ORGANISM      pan troglodytes
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE      1 (sites)
AUTHORS      Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
      Totoki,Y., Watanabe,H. and Sakaki,Y.
      BAC end sequences of library PTB
      Unpublished
      2 (bases 1 to 1030)
      Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
      Totoki,Y., Watanabe,H. and Sakaki,Y.
      Direct Submission
      Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
      and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
      1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
      (E-mail:chimpbes@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/)
      Tel:81-45-503-9111, Fax:81-45-503-9170)
      Clones are derived from the chimpanzee BAC library PTB This BAC end
      clones generated during the Rsd process and may have higher chance of
      clone tracking errors.
      PRIMERS
      Sequencing: -21M13
      LIBRARY
      Vector : pKS145
      R.Site 1 : SacI
      R.Site 2 : SacI.
      Location/Qualifiers
FEATURES
      source
      1..1030
      /organism="Pan troglodytes"
      /db_xref="taxon:9598"
      /clone="PTB-150D13.F"
      /sex="male"
      /cell_type="lymphoblast"
      /clone_lib="PTB chimpanzee Male BAC library"
BASE COUNT      239 a      354 c      207 g      198 t      32 others

Query Match      84.0%; Score 16.8; DB 12; Length 1030;
Best Local Similarity 90.0%; Pred. No. 3e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 9999gacatcgtcg9999g 20
      11111 11111 11111
Db      755 GGGGGGCGATCGTGGGGG 736

RESULT 12      1081 bp      mRNA      linear      EST 05-FEB-2002
BM470460      BM470460      AGENCOURT-6462980 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5533377
LOCUS      5', mRNA sequence.
DEFINITION      BM470460
ACCESSION      BM470460
VERSION      BM470460.1 GI:18519502
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1081)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
      National Institutes of Health, Mammalian Gene Collection (MGC)
      Unpublished (1999)
      Contact: Robert Strausberg, Ph.D.
      Email: cgabbs-remail.nih.gov
      Tissue Procurement: ARCC
      CDNA Library Preparation: Life Technologies, Inc.
      DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
      DNA Sequencing by: Agencourt Bioscience Corporation
      Clone distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12218 row: c column: 10

High quality sequence start: 17

High quality sequence stop: 671.

location/Qualifiers

1.1081

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:533377"

/clone_1lb="NIH_MGC_71"

/tissue_type="leiomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: uterus; Vector: pCMV-Sport6; Site:1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 2.1 kb.

236 a 335 c 304 g 205 t 1 others

BASE COUNT

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 1081;

Best Local Similarity 90.0%; Pred. No. 3e+03; 2; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1066 GGGGACGACGTCGCGGCG 1047

OY 1 ggggagcagtcgtcgggggg 20

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

BASE COUNT 203 a 321 c 323 g 216 t 38 others

ORIGIN

Query Match

Best Local Similarity 90.0%;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggggagcagtcgtcgggggg 20

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

BASE COUNT 255 a 289 c 361 g 215 t

ORIGIN

Query Match

Best Local Similarity 90.0%;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggggagcagtcgtcgggggg 20

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

BASE COUNT 1347 bp mRNA linear EST 05-FEB-2002

ORIGIN

Query Match

Best Local Similarity 90.0%;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggggagcagtcgtcgggggg 20

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

BASE COUNT 1347 bp mRNA linear EST 05-FEB-2002

ORIGIN

Query Match

Best Local Similarity 90.0%;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggggagcagtcgtcgggggg 20

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

BASE COUNT 1347 bp mRNA linear EST 05-FEB-2002

ORIGIN

Query Match

Best Local Similarity 90.0%;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggggagcagtcgtcgggggg 20

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

BASE COUNT 1347 bp mRNA linear EST 05-FEB-2002

ORIGIN

Query Match

Best Local Similarity 90.0%;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggggagcagtcgtcgggggg 20

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

BASE COUNT 1347 bp mRNA linear EST 05-FEB-2002

ORIGIN

Query Match

Best Local Similarity 90.0%;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggggagcagtcgtcgggggg 20

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1347)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC).
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LAM12322 row: b column: 24
 High quality sequence start: 99
 High quality sequence stop: 279.
 Location/Qualifiers
 1..1347
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5574071"
 /clone_lib="NIH_MGC_88"
 /tissue_type="duodenal adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: small intestine; Vector: pCMV-SPORT6;
 Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
 oligo-dT primed. Average insert size 1.767 kb. Library
 enriched for full-length clones and constructed by Life
 Technologies. Note: this is a NIH-MGC Library."
 BASE COUNT 299 a 548 c 285 g 205 t 10 others
 ORIGIN
 Query Match 84.0%; Score 16.8; DB 10; Length 1347;
 Best Local Similarity 90.0%; Pred. No. 3.1e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 99999acgacgtcggggg 20
 ||||| ||| |||||
 Db 576 GGGGACGCTCGGGGGG 557

Search completed: August 10, 2002, 02:11:05
 Job time: 13126 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2002, 22:36:34 ; Search time 2778.35 Seconds
(without alignments)
150.640 Million cell updates/sec

Title: US-09-672-126-7

Perfect score: 20
Sequence: 1 gggggagacatcgcggggg 20

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.ov:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htgo.inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	20	100.0	20	6	AX105109	Sequence
2	20	100.0	20	6	AX105234	Sequence
3	20	100.0	20	6	AX104805	Sequence
4	20	100.0	20	6	AX104806	Sequence
5	20	100.0	20	6	AX105117	Sequence
6	20	100.0	20	6	AX105118	Sequence
7	19	95.0	20	6	AX104767	Sequence
8	19	95.0	20	6	AX104883	Sequence
9	19	95.0	20	6	AX105137	Sequence
10	18.4	92.0	20	6	AX104803	Sequence
11	18.4	92.0	20	6	AX104804	Sequence
12	18.4	92.0	20	6	AX105115	Sequence
13	18.4	92.0	20	6	AX105116	Sequence
14	18	90.0	19	6	AX104879	Sequence
15	18	90.0	19	6	AX105134	Sequence
16	17.4	87.0	20	2575	AC096281	Sequence
17	17.4	87.0	20	2575	AC099298	Sequence
18	16.8	84.0	20	6	AX104799	Sequence
19	16.8	84.0	20	6	AX105114	Sequence
20	16.8	84.0	20	6	AF125967	Sequence
21	16.8	84.0	48872	2	AC099246	Sequence
22	16.8	84.0	122056	2	D90902	Sequence
23	16.8	84.0	183512	2	AC095258	Sequence
24	16.8	84.0	190850	2	AC094600	Sequence
25	16.8	84.0	315000	1	RME603644	Sequence
26	16.4	82.0	20	6	AX104864	Sequence
27	16.4	82.0	1776	8	AR041208	Sequence
28	16.4	82.0	1776	8	ASNRHBA	Sequence
29	16.4	82.0	145709	1	D90914	Sequence
30	16.4	82.0	4120	14	MFO238973	Sequence
31	16	80.0	195859	16	AF281817	Sequence
32	15.8	79.0	1887	1	AF058788	Sequence
33	15.8	79.0	2267	1	RHMAGTS	Sequence
34	15.8	79.0	10802	2	AE005764	Sequence
35	15.8	79.0	12956	2	AC109435	Sequence
36	15.8	79.0	33755	2	AC012855	Sequence
37	15.8	79.0	36615	2	AC101782	Sequence
38	15.8	79.0	48603	2	AC094658	Sequence
39	15.8	79.0	53190	2	AC102338	Sequence
40	15.8	79.0	61779	2	AC104888	Sequence
41	15.8	79.0	67312	2	AC106197	Sequence
42	15.8	79.0	85412	2	AC096236	Sequence
43	15.8	79.0	85538	2	AC093986	Sequence
44	15.8	79.0	87833	2	AC064805	Sequence
45	15.8	79.0	88065	9	AP002015	Sequence

ALIGNMENTS

RESULT 1
AX105109
LOCUS AX105109 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 7 from Patent WO0122990.
ACCESSION AX105109
VERSION AX105109.1 GI:13921259
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Interferon
JOURNAL Patent: WO 0122990-A 7 (05-APR-2001);
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"

misc_feature 1..2

misc_feature 3..14 /note="Backbone has phosphorothioate linkages."
misc_feature 15..19 /note="Backbone has phosphodiester linkages."
misc_feature 20 /note="Backbone has phosphorothioate linkages."
BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 9999gacgacgtcgcg999g 20
1 GGGGGACGATCGTCGGGGG 20

RESULT 2
AX105234 20 bp DNA linear PAT 30-APR-2001
LOCUS AX105234
DEFINITION Sequence 133 from Patent WO0122990.
ACCESSION AX105234
VERSION AX105234.1 GI:13921384
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
JOURNAL Interferon
Patent: WO 0122990-A 133 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)

FEATURES
source location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic oligonucleotide"

BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 9999gacgacgtcgcg999g 20
1 GGGGGACGATCGTCGGGGG 20

RESULT 3
AX104805 21 bp DNA linear PAT 30-APR-2001
LOCUS AX104805
DEFINITION Sequence 997 from Patent WO0122972.
ACCESSION AX104805
VERSION AX104805.1 GI:13921002
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 21)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 997 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)

FEATURES
source location/Qualifiers
1..21

/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 2 a 3 c 14 g 2 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 9999gacgacgtcgcg999g 20
2 GGGGGACGATCGTCGGGGG 21

RESULT 4
AX104806 21 bp DNA linear PAT 30-APR-2001
LOCUS AX104806
DEFINITION Sequence 998 from Patent WO0122972.
ACCESSION AX104806
VERSION AX104806.1 GI:13921003
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 21)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 998 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)

FEATURES
source location/Qualifiers
1..21
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic oligonucleotide"

BASE COUNT 2 a 3 c 14 g 2 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 9999gacgacgtcgcg999g 20
1 GGGGGACGATCGTCGGGGG 20

RESULT 5
AX105117 21 bp DNA linear PAT 30-APR-2001
LOCUS AX105117
DEFINITION Sequence 15 from Patent WO0122990.
ACCESSION AX105117
VERSION AX105117.1 GI:13921267
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 21)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
JOURNAL Interferon
Patent: WO 0122990-A 15 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)

FEATURES
source location/Qualifiers
1..21
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic oligonucleotide"

misc_feature 1..2
misc_feature 3..15
misc_feature /note="Backbone has phosphorothioate linkages."

misc_feature /note="Backbone has phosphodiester linkages."
16..20
misc_feature 21 /note="Backbone has phosphorothioate linkages."
BASE COUNT 2 a 3 c 14 g 2 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 99999acgacgtcgcggggg 20
|||||
Db 2 GGGGACGATCGTCGGGGG 21

RESULT 6 AX105118 21 bp DNA linear PAT 30-APR-2001
LOCUS AX105118 Sequence 16 from Patent WO0122990.

DEFINITION AX105118
ACCESSION AX105118
VERSION AX105118.1 GI:13921268
KEYWORDS

SOURCE synthetic construct.
ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 21)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.

TITLE Methods related to immunostimulatory nucleic acid-induced interferon

JOURNAL Patent: WO 0122990-A 16 05-APR-2001;

COLEY Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)

FEATURES Location/Qualifiers
1..21

source /organism="synthetic construct"
/db_xref="taxon:32630"

misc_feature /note="Synthetic Oligonucleotide"

1..2 /note="Backbone has phosphorothioate linkages."

3..15 /note="Backbone has phosphodiester linkages."

16..20 /note="Backbone has phosphorothioate linkages."

21 /note="Backbone has phosphodiester linkages."

BASE COUNT 2 a 3 c 14 g 2 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 99999acgacgtcgcggggg 20
|||||
Db 1 GGGGACGATCGTCGGGGG 20

RESULT 7 AX104767 19 bp DNA linear PAT 30-APR-2001
LOCUS AX104767 Sequence 959 from Patent WO0122972.

DEFINITION AX104767
ACCESSION AX104767
VERSION AX104767.1 GI:13920964
KEYWORDS

SOURCE synthetic construct.
ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 19)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.

TITLE Methods related to immunostimulatory nucleic acid-induced interferon

JOURNAL Patent: WO 0122990-A 16 05-APR-2001;

COLEY Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)

FEATURES Location/Qualifiers
1..20

TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 959 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical GmbH (DE)

FEATURES Location/Qualifiers
1..19
source /organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 2 a 3 c 12 g 2 t
ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 99999acgacgtcgcggggg 19
|||||
Db 1 GGGGACGATCGTCGGGGG 19

RESULT 8 AX104883 20 bp DNA linear PAT 30-APR-2001
LOCUS AX104883 Sequence 1075 from Patent WO0122972.

DEFINITION AX104883
ACCESSION AX104883
VERSION AX104883.1 GI:13921080
KEYWORDS

SOURCE synthetic construct.
ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.

TITLE Immunostimulatory nucleic acids

JOURNAL Patent: WO 0122972-A 1075 05-APR-2001;

UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical GmbH (DE)

FEATURES Location/Qualifiers
1..20

source /organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 99999acgacgtcgcggggg 20
|||||
Db 1 GGGGACGATCGTCGGGGG 19

RESULT 9 AX105137 20 bp DNA linear PAT 30-APR-2001
LOCUS AX105137 Sequence 35 from Patent WO0122990.

DEFINITION AX105137
ACCESSION AX105137
VERSION AX105137.1 GI:13921287
KEYWORDS

SOURCE synthetic construct.
ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 20)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.

TITLE Methods related to immunostimulatory nucleic acid-induced interferon

JOURNAL Patent: WO 0122990-A 35 05-APR-2001;

COLEY Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)

FEATURES Location/Qualifiers
1..20

source

misc_feature /organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 1. .2
/note="Synthetic Oligonucleotide"
misc_feature 3. .14
/note="Backbone has phosphorothioate linkages."
misc_feature 15. .19
/note="Backbone has phosphorothioate linkages."
misc_feature 20
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 99ggacgacgtcgcggggg 20
|||||
Db 1 GGGGACGATCGTCGGGGG 19

RESULT 10
AX104803 20 bp DNA linear PAT 30-APR-2001
LOCUS
DEFINITION Sequence 995 from Patent WO0122972.
ACCESSION AX104803
VERSION AX104803.1 GI:13921000
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 20)
AUTHORS Kriegl,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 995 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US); Coley Pharmaceutical
GmbH (DE)

FEATURES
source location/Qualifiers
1. .20
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 2 a 2 c 13 g 3 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 20;
Best Local Similarity 95.0%; Pred. No. 8.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 99gggacgacgtcgcggggg 20
|||||
Db 1 GGGGACGATCGTCGGGGG 20

RESULT 11
AX104804 20 bp DNA linear PAT 30-APR-2001
LOCUS
DEFINITION Sequence 996 from Patent WO0122972.
ACCESSION AX104804
VERSION AX104804.1 GI:13921001
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 20)
AUTHORS Kriegl,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 996 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US); Coley Pharmaceutical
GmbH (DE)

FEATURES
source location/Qualifiers
1. .20
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 3 a 3 c 12 g 2 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 20;
Best Local Similarity 95.0%; Pred. No. 8.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 99gggacgacgtcgcggggg 20
|||||
Db 1 GGGGACGATCGTCGGGGG 20

RESULT 12
AX105115 20 bp DNA linear PAT 30-APR-2001
LOCUS
DEFINITION Sequence 13 from Patent WO0122990.
ACCESSION AX105115
VERSION AX105115.1 GI:13921265
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 20)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Kriegl,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
Interferon
JOURNAL Patent: WO 0122990-A 13 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US); UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)

FEATURES
source location/Qualifiers
1. .20
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 1. .2
/note="Backbone has phosphorothioate linkages."
misc_feature 3. .15
/note="Backbone has phosphodiester linkages."
misc_feature 16. .19
/note="Backbone has phosphorothioate linkages."
misc_feature 20
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 2 c 13 g 3 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 20;
Best Local Similarity 95.0%; Pred. No. 8.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 99gggacgacgtcgcggggg 20
|||||
Db 1 GGGGACGATCGTCGGGGG 20

RESULT 13
AX105116 20 bp DNA linear PAT 30-APR-2001
LOCUS
DEFINITION Sequence 14 from Patent WO0122990.
ACCESSION AX105116
VERSION AX105116.1 GI:13921266
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 20)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Kriegl,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced

JOURNAL Interferon
Patent: WO 0122990-A 14 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)

FEATURES

source

Location/Qualifiers

1. .20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"

misc_feature 1. .2
/note="Backbone has phosphorothioate linkages."

misc_feature 3. .15
/note="Backbone has phosphodiester linkages."

misc_feature 16. .19
/note="Backbone has phosphorothioate linkages."

misc_feature 20
/note="Backbone has phosphodiester linkages."

BASE COUNT 3 a 3 c 12 g 2 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 20;

Best Local Similarity 95.0%; Pred. No. 8.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gggagcagtcgtcgggggg 20
|||||
Db 1 GGGAGCAGTCGTCGGGGG 20

RESULT 14

AX104879 19 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1071 from Patent WO0122972.
ACCESSION AX104879
VERSION AX104879.1 GI:13921076

KEYWORDS

SOURCE synthetic construct.
ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 19)
Krieg, A.M., Schetter, C. and Volmer, J.C.

JOURNAL Immunostimulatory nucleic acids
Patent: WO 0122972-A 1071 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)

FEATURES

source

Location/Qualifiers

1. .19
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 2 a 3 c 12 g 2 t
ORIGIN

Query Match 90.0%; Score 18; DB 6; Length 19;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 gggagcagtcgtcgggggg 20
|||||
Db 1 GGGAGCAGTCGTCGGGGG 18

RESULT 15

AX105134 19 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 32 from Patent WO0122990.
ACCESSION AX105134
VERSION AX105134.1 GI:13921284

KEYWORDS

SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.

REFERENCE

1 (bases 1 to 19)
Hartmann, G.D., Bratzler, R.L. and Krieg, A.U.

JOURNAL Methods related to immunostimulatory nucleic acid-induced
Interferon
Patent: WO 0122990-A 32 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)

FEATURES

source

Location/Qualifiers

1. .19
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"

misc_feature 1. .2
/note="Backbone has phosphorothioate linkages."

misc_feature 3. .13
/note="Backbone has phosphodiester linkages."

misc_feature 14. .18
/note="Backbone has phosphorothioate linkages."

BASE COUNT 2 a 3 c 12 g 2 t
ORIGIN

Query Match 90.0%; Score 18; DB 6; Length 19;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 gggagcagtcgtcgggggg 20
|||||
Db 1 GGGAGCAGTCGTCGGGGG 18

Search completed: August 10, 2002, 02:57:36
Job time: 15662 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:21:45 ; Search time 1145.36 Seconds
(without alignments)
32.978 Million cell updates/sec

Title: US-09-672-126-9

Perfect score: 22
Sequence: 1 gggggacatcgtcgggggg 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	AAF98739	Human IFN-alpha 1m
2	22	100.0	22	AAF99783	Immunostimulatory
3	18.8	85.5	22	AAF98740	Human IFN-alpha 1m
4	18.8	85.5	22	AAF98741	Human IFN-alpha 1m
5	18.8	85.5	22	AAF99784	Immunostimulatory
6	18.8	85.5	22	AAF99785	Immunostimulatory
7	17.8	80.9	308	AAC03250	Human secreted pro
8	17.8	80.9	400	ABA08414	Human secreted pro
9	17.8	80.9	726	AAH04520	Human CDNA clone (

10	17.8	80.9	1014	22	AAS41225	CDNA encoding nove
11	17.8	80.9	1401	22	AAI60544	Human polynucleoti
12	17.8	80.9	1904	22	AAC75713	Human ORF2186
13	17.8	80.9	1905	22	AAI58758	Human polynucleoti
14	17.8	80.9	5796	22	AAS42029	Genomic sequence #
15	17.4	79.1	8718	24	ABU33273	Human immune syste
16	16.2	73.6	21	22	AAF98767	Human IFN-alpha 1m
17	16.2	73.6	21	22	AAF98767	Immunostimulatory
18	16.2	73.6	369	21	AAC01688	Human secreted pro
19	16.2	73.6	784	22	AAH06823	Human CDNA clone (
20	16.2	73.6	907	21	AAC76631	Human ORF2186
21	16.2	73.6	1282	22	AAI58615	Human polynucleoti
22	16.2	73.6	1419	22	AAF61001	P. putida K12440-a
23	16.2	73.6	1470	22	AAI60401	Human polynucleoti
24	16.2	73.6	1929	17	AAIT10954	Chicken adenovirus
25	16.2	73.6	2148	15	AAO73222	Bovine parathyroid
26	16.2	73.6	2642	22	AAH14538	Human CDNA sequenc
27	16.2	73.6	5275	19	AAV26962	Bovine parathyroid
28	16.2	73.6	5275	19	AAV26962	Bovine parathyroid
29	16.2	73.6	5275	20	AAV25053	Bovine parathyroid
30	16.2	73.6	5275	20	AAV82483	Bovine parathyroid
31	16.2	73.6	5275	21	AAE89296	Bovine parathyroid
32	16.2	73.6	5275	24	AAI72120	CDNA encoding BOPC
33	16.2	73.6	9021	22	AAI46326	Tumour suppressor
34	16.2	73.6	16235	22	AAK86192	Human immune/haema
35	16.2	73.6	23128	23	AAI59552	Proionibacterium
36	16.2	73.6	38186	20	AAZ32028	Human MERT1 relate
37	16.2	73.6	38186	22	AAC90085	Human CDNA cion
38	16.2	73.6	43804	18	AAE86575	Chicken embryo let
39	16.2	73.6	43804	20	AAE86590	Complete genome se
40	16.2	73.6	44018	22	AAE82392	Human adenovirus C
41	15.8	71.8	750	21	AAC77592	Human ORF23147
42	15.8	71.8	1621	22	AAC60228	Human hydroxylase-1
43	15.8	71.8	1636	22	AAI60336	Human polynucleoti
44	15.8	71.8	1650	22	AAI58550	Human polynucleoti
45	15.8	71.8	6048	24	ABL34030	Human immune syste

ALIGNMENTS

RESULT 1	
ID AAF98739	AAF98739 standard; DNA; 22 BP.
XX AAF98739;	
AC AAF98739;	
XX AAF98739;	
DT 11-JUN-2001	(first entry)
XX 11-JUN-2001	
DE Human IFN-alpha	immunostimulatory nucleic acid SEQ ID NO: 9.
XX Immunostimulatory	nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection;	phosphorothioate backbone; palindrome; cancer; ds.
KW	
OS Synthetic.	
XX	
PH key	Location/Qualifiers
FT modified_base	1..2
FT	/tag= a
FT	/mod_base= "OTHER"
FT	/note= "phosphorothioate linkage"
FT modified_base	17..21
FT	/tag= b
FT	/mod_base= "OTHER"
FT	/note= "phosphorothioate linkage"
PD WO2001222990-A2.	
XX 05-APR-2001.	
XX	
PF 27-SEP-2000;	2000WO-US26527.
XX	
PR 27-SEP-1999;	99US-0156147.

XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.
XX Hartmann G, Bratzler RL, Krieg A;
DR WPI; 2001-290487/30.
XX
PT Improving the efficacy of treatments involving the administration of
PT Interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
PS Claim 201; Page 103; 168pp; English.
XX
CC The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX
SQ Sequence 22 BP; 3 A; 3 C; 13 G; 3 T; 0 other;

Query Match 100.0%; Score 22; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 99999acgatacgtcggggg 22
1 99999acgatacgtcggggg 22

Db 1 99999acgatacgtcggggg 22

RESULT 2
AAF99783
ID AAF99783 standard; DNA; 22 BP.
AC AAF99783;
XX
DT 12-JUN-2001 (first entry)
XX
DE Immunostimulatory nucleic acid #899.
XX
KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss;
XX
OS Synthetic.
XX
PN WO200122972-A2.
XX
PD 05-APR-2001.
XX
PE 25-SEP-2000; 2000WO-US26383.
XX
PF 25-SEP-1999; 99US-0156113.
XX
PR 27-SEP-1999; 99US-0156135.
XX
PR 23-AUG-2000; 2000US-0227436.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX
DR WPI; 2001-273485/28.
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 57; 338pp; English.

XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC hemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC T_H2 to a T_H1 immune response and to activate immune cells.
XX
SQ Note: the present sequence may have a phosphorothioate backbone.
SQ Sequence 22 BP; 3 A; 3 C; 13 G; 3 T; 0 other;

Query Match 100.0%; Score 22; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 99999acgatacgtcggggg 22
1 99999acgatacgtcggggg 22

Db 1 99999acgatacgtcggggg 22

RESULT 3
AAF98740
ID AAF98740 standard; DNA; 22 BP.
AC AAF98740;
XX
DT 11-JUN-2001 (first entry)
XX
DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 10.
XX
KW Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..2
FT /*tag= a
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
FT modified_base 17..21
FT /*tag= b
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
XX
PN WO200122990-A2.
XX
PD 05-APR-2001.
XX
PE 27-SEP-2000; 2000WO-US26527.
XX
PF 27-SEP-1999; 99US-0156147.
XX
PR 27-SEP-1999; 99US-0156147.
XX
PR (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.
XX
PI Hartmann G, Bratzler RL, Krieg A;
XX
DR WPI; 2001-290487/30.
XX
PT Improving the efficacy of treatments involving the administration of
PT Interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
PS Claim 201; Page 103; 168pp; English.

XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering the
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.

XX
SQ Sequence 22 BP; 2 A; 4 C; 14 G; 2 T; 0 other;

Query Match 85.5%; Score 18.8; DB 22; Length 22;
Best Local Similarity 90.9%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 9999gacgacatcgtcg9999 22
||||||| |||||||
Db 1 9999gacgacgtcgtcg9999 22

RESULT 4
AAF98741
ID AAF98741 standard; DNA; 22 BP.

AC AAF98741;
DT 11-JUN-2001 (first entry)

DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 11.

KM Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.

XX Synthetic.

OS
XX Key location/Qualifiers

FT modified_base 1..2

FT /*tag= a

FT /mod_base= "OTHER"

FT /note= "phosphorothioate linkage"

FT modified_base 17..21

FT /*tag= b

FT /mod_base= "OTHER"

FT /note= "phosphorothioate linkage"

FT WO200122990-A2.

PD 05-APR-2001.

PF 27-SEP-2000; 2000WO-US26527.

PR 27-SEP-1999; 990S-0156147.

PA (COLE-) COLEY PHARM GROUP INC.

PI (IOWA) UNIV IOWA RES FOUND.

PI Hartmann G, Bratzler RL, Krieg A;

PI WPI; 2001-290487/30.

DR Improving the efficacy of treatments involving the administration of

PT interferon-alpha by co-administering an isolated immunostimulatory

PT nucleic acid -

PS Claim 201; Page 103; 168bp; English.

CC The present invention describes an improvement to a method requiring the

CC administration of interferon alpha (IFN-alpha), involving administering

CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of

CC such nucleic acids are also provided. These may comprise oligonucleotides

CC with phosphorothioate backbones, palindromes, or G-rich sequences. The

CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.

XX
SQ Sequence 22 BP; 2 A; 4 C; 14 G; 2 T; 0 other;

Query Match 85.5%; Score 18.8; DB 22; Length 22;
Best Local Similarity 90.9%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 9999gacgacatcgtcg9999 22
||||||| |||||||
Db 1 9999gacgacgtcgtcg9999 22

RESULT 5
AAF9784
ID AAF9784 standard; DNA; 22 BP.

AC AAF9784;

DT 12-JUN-2001 (first entry)

DE Immunostimulatory nucleic acid #900.

KM Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;

KW immunostimulatory; tumour; viral infection; bacterial infection;

KW fungal infection; parasitic infection; cancer; asthma;

KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.

XX Synthetic.

OS
XX WO200122972-A2.

PD 05-APR-2001.

PF 25-SEP-2000; 2000WO-US26383.

PR 25-SEP-1999; 990S-0156113.

PR 27-SEP-1999; 990S-0156135.

PR 23-AUG-2000; 2000US-0227436.

XX (IOWA) UNIV IOWA RES FOUND.

PA (COLE-) COLEY PHARM GMBH.

PI Krieg AM, Schetter C, Volmer J;

PI WPI; 2001-273485/28.

DR Vaccinating against tumors, infectious diseases, allergies and asthma

PT using immunostimulatory Py-rich and TG nucleic acids -

PS Claim 101; Page 57; 338bp; English.

CC The present invention relates to a method for stimulating an immune

CC response. The method comprises administering an immunostimulatory nucleic

CC acid to a non-rodent subject in sufficient quantity to stimulate an

CC immune response. The present sequence is one such immunostimulatory

CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich

CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects

CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae

CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,

CC haemophilus, campylobacter, clostridium, Escherichia coli and/or

CC staphylococcus), fungal antigens and/or parasitic antigens. The method is

CC also useful for preventing cancer, asthma, infectious disease, allergy or

CC immune deficiency. The present sequence can also be used to redirect a

CC Th2 to a Th1 immune response and to activate immune cells.

CC Note: the present sequence may have a phosphorothioate backbone.

XX
SQ Sequence 22 BP; 2 A; 4 C; 14 G; 2 T; 0 other;

Query Match 85.5%; Score 18.8; DB 22; Length 22;
 Best Local Similarity 90.9%; Pred. No. 15;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggagcatatcgtcggggg 22
 |||||
 Db 1 gggggagcagctcgtcggggg 22

RESULT 6

AAF99785
 ID AAF99785 standard; DNA; 22 BP.

AC AAF99785;

DT 12-JUN-2001 (first entry)

DE Immunostimulatory nucleic acid #901.

KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;

KM immunostimulatory; tumour; viral infection; bacterial infection;

KW fungal infection; parasitic infection; cancer; asthma;

KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.

OS Synthetic.

PN W0200122972-A2.

PD 05-APR-2001.

PF 25-SEP-2000; 2000WO-US26383.

PR 25-SEP-1999; 99US-0156113.

PR 27-SEP-1999; 99US-0156135.

PR 23-AUG-2000; 2000US-0227436.

PA (IOWA) UNIV IOWA RES FOUND.

PA (COLE-) COLEY PHARM GMBH.

PI Kriegl AM, Schetter C, Vollmer J;

DR WPI; 2001-273485/28.

PT Vaccinating against tumors, infectious diseases, allergies and asthma

PS using immunostimulatory Py-rich and TG nucleic acids -

PS Claim 101; Page 57; 338pp; English.

XX The present invention relates to a method for stimulating an immune

CC response. The method comprises administering an immunostimulatory nucleic

CC acid to a non-rodent subject in sufficient quantity to stimulate an

CC immune response. The present sequence is one such immunostimulatory

CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich

CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects

CC against tumor antigens, viral antigens (e.g. herpesviridae, retroviridae

CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,

CC hemophilus, campylobacter, clostridium, Escherichia coli and/or

CC stephylococcus), fungal antigens and/or parasitic antigens. The method is

CC also useful for preventing cancer, asthma, infectious disease, allergy or

CC immune deficiency. The present sequence can also be used to redirect a

CC T12 to a Th1 immune response and to activate immune cells.

Note: the present sequence may have a phosphorothioate backbone.

Sequence 22 BP; 2 A; 4 C; 14 G; 2 T; 0 other;

Query Match 85.5%; Score 18.8; DB 22; Length 22;

Best Local Similarity 90.9%; Pred. No. 15;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggagcatatcgtcggggg 22
 |||||
 Db 1 gggggagcagctcgtcggggg 22

RESULT 7

AAC03250
 ID AAC03250 standard; CDNA; 308 BP.

AC AAC03250;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 3248.

KW Human; 5' EST; expressed sequence tag; secreted protein; CDNA isolation;

KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GENET) GENSET.

PI Dumas Milne Edwards J, Duclet A, Giordano J;

DR WPI; 2000-500381/45.

DR P-PSDB; AAG03244.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for

PT obtaining cdnas and genomic DNAs that correspond to 5'ESTs and for

PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 1; SEQ ID 3248; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from

CC mRNAs encoding secreted proteins. An ORF has been identified within the

CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs

CC derived from 30 different tissues. EST sequences usually correspond

CC mainly to the 3' untranslated region (UTR) of the mRNA because they are

CC often obtained from oligo-dT primed CDNA libraries. Such ESTs are not

CC well suited for isolating CDNA sequences derived from the 5' ends of

CC mRNAs and even in those cases where longer CDNA sequences have been

CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from

CC mRNAs with intact 5' ends and can therefore be used to obtain full length

CC CDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,

CC gene therapy and chromosome mapping procedures. They are used to obtain

CC upstream regulatory sequences and to design expression and secretion

CC vectors.

Sequence 308 BP; 68 A; 75 C; 105 G; 60 T; 0 other;

Query Match 80.9%; Score 17.8; DB 21; Length 308;

Best Local Similarity 90.5%; Pred. No. 44;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggagcatatcgtcggggg 21
 |||||
 Db 9 gggggagcagctcgtcgtg 29

RESULT 8

ABA08414
 ID ABA08414 standard; CDNA; 400 BP.

AC ABA08414;

DT 11-JAN-2002 (first entry)

CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 726 BP; 152 A; 148 C; 247 G; 176 T; 3 other;

Query Match 80.9%; Score 17.8; DB 22; Length 726;
Best Local Similarity 90.5%; Pred. No. 44;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gggggagcagatcgctcgggg 21
||||||| ||||| ||
Db 8 gggggagcagatcgctcggtg 28

RESULT 10
AAS41225
ID AAS41225 standard; cDNA; 1014 BP.
XX
AC AAS41225;
XX
DT 17-DEC-2001 (first entry)
XX
DE cDNA encoding novel human enzyme polypeptide #441;
XX
KW Human: oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
XX
OS Homo sapiens.
XX
PN WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250191.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI: 2001-465566/50.

P-PSDB: AAU23335.

Novel polypeptides and polynucleotides useful for diagnosing,
preventing, treating neural, immune system, muscular, reproductive,
pulmonary, cardiovascular, renal, proliferative disorders and cancerous
diseases -

Claim 4: SEQ ID NO 451; 1180bp; English.

The present invention relates to the isolation of novel human enzyme
polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
encoding them. The enzyme polypeptides of the invention may comprise the
functional classes of oxidoreductases, transferases, hydrolases, lyases,
isomerases or ligases. The sequences of the invention are useful in the
diagnosis, treatment, prevention and/or prognosis of a wide range of
disorders including hyperproliferative disorders (e.g. cancer),
immunodeficiency disorders (e.g. AIDS) autoimmune disorders
(e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
metabolic disorders (e.g. phenylketonuria), inflammatory disorders
(e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
blood-related disorders (e.g. haemophilia), reproductive disorders
(e.g. infertility) and infectious disorders (e.g. influenza). The
polynucleotides of the invention can also be used in gene therapy.
AAU40788-AAU41684 represent cDNA sequences encoding for the novel human
enzyme polypeptides of the invention.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1014 BP; 222 A; 254 C; 335 G; 199 T; 4 other;

Query Match

80.9%; Score 17.8; DB 22; Length 1014;

Best Local Similarity 90.5%; Pred. No. 44;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gggggacagatctgcggagg 21
|||||
Db 45 gggggacagattcgcggtgg 65

RESULT 11

AAI60544

ID AAI60544 standard; cDNA; 1401 BP.

AC AAI60544;

XX XX

DT 22-OCT-2001 (first entry)

XX XX

DE Human polynucleotide SEQ ID NO 4533.

XX XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokine; thrombolytic; drug screening; arthritis; inflammation;

KW Leukemia; ss.

XX XX

OS Homo sapiens.

XX XX

PN WO200153312-A1.

XX XX

PD 26-JUL-2001.

XX XX

PF 26-DEC-2000; 2000WO-US34263.

XX XX

PR 21-JAN-2000; 2000US-0486725.

XX XX

PR 25-APR-2000; 2000US-0552317.

XX XX

PR 09-JUL-2000; 2000US-0598042.

XX XX

PR 19-JUL-2000; 2000US-0620312.

XX XX

PR 03-AUG-2000; 2000US-0653450.

XX XX

PR 14-SEP-2000; 2000US-0662191.

XX XX

PR 19-OCT-2000; 2000US-0693036.

XX XX

PR 29-NOV-2000; 2000US-0727344.

XX XX

PA (HYSE-) HYSEQ INC.

XX XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX XX

DR WPI: 2001-442253/47.

XX XX

DR P-PSDB: AAM41388.

XX XX

PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX XX

PS Claim 1: SEQ ID NO 4533; 10078bp; English.

XX XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic, and

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localized neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC actinin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

SQ Sequence 1401 BP; 293 A; 362 C; 457 G; 289 T; 0 other;
 Query Match 80.9%; Score 17.8; DB 22; Length 1401;
 Best Local Similarity 90.5%; Pred. No. 45;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 9999gacatcgtcgg99 21
 ||||||| |||||||
 Db 64 9999gacattcgtcgtg 84
 RESULT 12
 AAC75713
 ID AAC75713 standard; cDNA; 1904 BP.
 XX AAC75713;
 AC
 XX 08-FEB-2001 (first entry)
 DT
 XX
 DE Human ORFX ORF1268 polynucleotide sequence SEQ ID NO:2535.
 XX
 KM Human: open reading frame: ORFX; detection: cytosstatic; hepatotropic;
 KM vulnery; antiparKinsonian; neurotropic; neuroprotective;
 KM anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KM hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KM antiviral; antibacterial; antifungal; antirheumatic; antihypertoid;
 KM antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KM cholesterol ester storage; systemic lupus erythematosus; infection;
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KM allergy; aplastic anaemia; nocturnal hemoglobinuria; burn; wound;
 KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KM thrombosis; contraceptive; SS.
 KM
 XX Homo sapiens.
 OS
 XX WO200058473-A2.
 PN
 XX 05-OCT-2000.
 PD
 XX 31-MAR-2000; 2000MO-US08621.
 PF
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR P-PDB; AAB41504.
 XX
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 5; Page 1816-1818; 5507pp; English.
 PS
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosstatic; hepatotropic; vulnery;
 CC antiparKinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antihypertoid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating

CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal hemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX
 SQ Sequence 1904 BP; 424 A; 504 C; 586 G; 389 T; 1 other;
 Query Match 80.9%; Score 17.8; DB 21; Length 1904;
 Best Local Similarity 90.5%; Pred. No. 45;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 9999gacatcgtcgg99 21
 ||||||| |||||||
 Db 49 9999gacattcgtcgtg 69
 RESULT 13
 AAI58758
 ID AAI58758 standard; cDNA; 1905 BP.
 XX AAI58758;
 AC
 XX 22-OCT-2001 (first entry)
 DT
 XX
 DE Human polynucleotide SEQ ID NO 961.
 XX
 KM Human: neurotropic; immunosuppressant; cytosstatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KM leukemia; ss.
 KM
 XX Homo sapiens.
 OS
 XX WO200153312-A1.
 PN
 XX 26-JUL-2001.
 PD
 XX 26-DEC-2000; 2000MO-US34263.
 PF
 XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSO INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR P-PDB; AAM39602.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 961; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
CC
XX
SQ Sequence 1905 BP; 417 A; 508 C; 587 G; 393 T; 0 other;

Query Match 80.9%; Score 17.8; DB 22; Length 1905;
Best Local Similarity 90.5%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gggggacgatctgcggggg 21
|||||
DB 65 gggggacgatctgcggggg 85

RESULT 14
AAS42029
ID AAS42029 standard; DNA; 5796 BP.
XX
XX AAS42029;
AC
XX
DT 17-DEC-2001 (first entry)
XX
XX Genomic sequence #345 encoding novel human enzyme polypeptide.
DE
XX
XX Human: oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200155301-A2.
PN
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01239.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0188874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225577.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228282.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI: 2001-465566/50.

The present invention relates to the isolation of novel human enzyme polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of diseases including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. influenza). The polynucleotides of the invention can also be used in gene therapy. AAU1685-AAU2192 represent DNA sequences encoding for the novel human enzyme polypeptides of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 5796 BP; 1304 A; 1423 C; 1707 G; 1362 T; 0 other;

Query Match	80.98;	Score 17.8;	DB 22;	Length 5796;
Best Local Similarity	90.58;	Pred. No. 45;		
Matches 19;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY 1 gggggacgatatcgtcgggg 21
    |||||
Db 53 gggggacgatttcgtcgtg 73
```

RESULT 15
APR 22 2007

ID ABL33273 standard; DNA; 8718 BP.

AC ABL33273;

DT 26-MAR-2002 (first entry)

Human immune system associated gene SEQ ID NO: 1246.

Human; immune system disease; cytosine methylation; antiasthmatic;

KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;

KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia

KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW

XXXXXX

XX
XX
PN

XX 03-TAM-3003
PD

XX 02-III-2001: 2001WA-FB07537
 PE

XX 30-TTN-2000: 2000DF-1032529
DP

PR 01-SEP-2000; 2000DE-1043826.
YY

PA (EP1G-) EPIGENOMICS AG.
XX

PI Olek A, Prepenbrock C, Berlin K, XY

WP1; 2002-130909/L7.

PT Nucleic acid comprising fragment of chemically modified gene, useful for diagnosis and treatment of diseases associated with abnormal

cytosine methylation

PS Claim 1; SEQ ID NO 124b; 32pp + Sequence Listing; German.
XX

The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences

CC including eye diseases such as retinopathy, neovascular glaucoma and

CC leukæmia. Alzheimer's disease. AIDS. neurofibromatosis
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid

CC diseases The present sequence is a gene of the invention

Sequence 8718 BP: 2225 A: 183 C: 2124 G: 4186 T: 0 Other: 0

Query Match	79.18;	Score 17.4;	DB 24;	Length 8718;
Best Local Similarity	94.78;	Pred No 69.		

Matches	18;	Conservative	0;	Mismatches	1;	Indels	0;	Caps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

OY 2 999gacatacgtcg99 20
||| |||||
Db 3302 999gacatacgtcg99 3320

Search completed: August 10, 2002, 03:21:46
Job time: 13677 sec

THIS PAGE BLANK (USPTO)


```

/sex="male"
/cell_type="lymphoblast"
BASE COUNT      166 a 238 c 101 g 178 t 6 others
ORIGIN

Query Match      85.5%; Score 18.8; DB 12; Length 689;
Best Local Similarity 90.9%; Pred. No. 5.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 999ggacgatactgcggggg 22
    |||||
Db 482 GGGGAGCATATTGCGGGG 461

RESULT 2
LOCUS      AUI05748      50 bp      mRNA      linear      EST 30-AUG-2001
DEFINITION AUI05748 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
ACCESSION  AUI05748
VERSION     AUI05748
KEYWORDS    AUI05748.1 GI:13555269
SOURCE      EST.
ORGANISM    human.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 50)
REFERENCE   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
AUTHORS    'H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            'Y., Nakamura,Y., Suyama,A. and Sugano,S.
            Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
            EMBO Rep. 2 (5), 388-393 (2001)
TITLE       JOURNAL
            MEDLINE
            21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Medical Science, University of Tokyo
            Institute of Medical Science, Minatoku, Tokyo 108-8639, Japan
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            'S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES
    source
        1..50
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="CAS01601"
        /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT      6 a 10 c 18 g 16 t
ORIGIN

Query Match      80.9%; Score 17.8; DB 9; Length 50;
Best Local Similarity 90.5%; Pred. No. 8.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 999ggacgatactgcggggg 21
    |||||
Db 10 GGGGAGCATTTGCGGTGG 30

RESULT 3
LOCUS      AU099475      300 bp      mRNA      linear      EST 05-APR-2001
DEFINITION AU099475 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HS105826 similar to Homo sapiens cDNA clone:DKFZP564C103, mRNA
            sequence.
ACCESSION  AU099475
VERSION     AU099475.1 GI:13550604
KEYWORDS    EST.
SOURCE      human.

```

```

ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 300)
REFERENCE   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
AUTHORS    'H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            'K., Suyama,A. and Sugano,S.
            In silico mapping of the 5'-ends of human mRNAs using full-length
            oligo-capping method
            Oligo-capping method
            Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            'S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES
    source
        1..300
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="HS105826"
        /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT      63 a 70 c 105 g 62 t
ORIGIN

Query Match      80.9%; Score 17.8; DB 9; Length 300;
Best Local Similarity 90.5%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 999ggacgatactgcggggg 21
    |||||
Db 52 GGGGAGCATTTGCGGTGG 72

RESULT 4
LOCUS      R54355      519 bp      mRNA      linear      EST 18-MAY-1995
DEFINITION y974g11.r1 Soares infant brain IN1B Homo sapiens cDNA clone
ACCESSION  R54355
VERSION     R54355
KEYWORDS    R54355.1 GI:816257
SOURCE      EST.
ORGANISM    human.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 519)
REFERENCE   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
AUTHORS    'M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            'R., Williamson,A., Woldmann,P. and Wilson,R.
            The WashU-Merck EST Project
            Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 2011
            High quality sequence stops: 393 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2011 Std Error: 0.00
            Seq primer: M13Rpi
            High quality sequence stop: 393.
            Location/Qualifiers
            1..519
            /organism="Homo sapiens"

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCAP613"
/clone_11b="Pediatric acute myelogenous leukemia cell (FAB
M1) Baylor-HGSC project-TCOA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="myeloid cell"
/dev_stage="pediatric 6 years"
/lab_host="DH10b"
/notice="Vector: lambda psb, Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer (5'GGAGACTCGAGCGCCGCAAGAGAG(T)VN
3'; V=A,C,G; N=A,C,G,T) and then dg tailed. Second strand
was primed with a BamHI-dC primer
(5'AGAGAGCTCGAGATCGCGCCGCAATTAATATAT(C) 3').

```

Query Match	80.9%;	Score 17.8;	DB 10;	Length 608;
Best Local Similarity	90.5%;	Pred. No. 1.3+03;		
Matches 19;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
1	gggggacgatactgcgcgggg	21		

Db	24	GCGGACGATTCGTGGTGG	44
RESULT	7		
LOCUS	BE264523		
DEFINITION	601192672p1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3536872 5'	mRNA	EST 13-JUL-2000
ACCESSION	BE264523		
VERSION	BE264523.1	GI:9138080	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Eutheria; Euthera; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Plate: LCM219 row: o column: 17 High quality sequence stop: 586. Location/Qualifiers 1. 616 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3536872" /clone_lib="NIH_MGC_7" /tissue_type="small cell carcinoma" /cell_line="MGC3" /lab_host="DH10B (phage-resistant)" /note="Organ: lung; Vector: pGB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGCAG(C). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."		
BASE COUNT	131 a 147 c 197 g 141 t		
ORIGIN			
Query Match	80.9%; Score 17.8; DB 10; Length 616; Best Local Similarity 90.5%; Pred. No. 1.3e+03; Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
OY	1_ggggacgatctgcggagg 21 		
Db	15_ggggacgatttcgtcggtgg 35		
RESULT	8		
LOCUS	BE265770		
DEFINITION	601194543p1 NIH_MGC_7 Homo sapiens CDNA clone IMAGE:3536186 5'	mRNA	linear EST 13-JUL-2000
ACCESSION	BE265770		
VERSION	BE265770.1	GI:9139251	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Plate: LCM223 row: f column: 11		

```

FEATURES                                High quality sequence stop: 234 .
SOURCE                                  Location/Qualifiers
    1. 630
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:3538186"
        /clone_1lb="NIH-MGC-7"
        /tissue.type="small cell carcinoma"
        /cell_line="MGC3"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: Lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT                             137 a          157 c          173 g          163 t

ORIGIN

Query Match                           80.9%; Score 17.8; DB 10; Length 630;
Best Local Similarity                 90.5%; Pred.No.1.3e+03;
Matches                               19; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

QY      1 999gacgatactcgcgg99g 21
         ||||||| ||||| ||
Db       15 ggcgacgatattcgtcggtgcg 35

RESULT      9
Bg761603   Bg761603           664 bp     mRNA             linear      EST 15-MAY-2001
DEFINITION 60271784/F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4841650 5',
LOCUS       mRNA sequence.
ACCESSION   Bg761603
VERSION     Bg761603.1 GI:14072256
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 664)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health; Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs@email.nih.gov
            Tissue Procurement: ATCC/DCTP/DRP
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.lnl.gov
            Plate: LCM1674 row: m column: 11
            High quality sequence stop: 664.
            Location/Qualifiers
                1..664
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:4841650"
                    /clone_1lb="NIH-MGC-49"
                    /tissue.type="melanotic melanoma, high MDR (cell line)"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript

```


II RT (Life Technologies). Note: this is a NIH_MGC library. |"

BASE COUNT 158 a 156 c 216 g 134 t

ORIGIN

Query Match 80.9%; Score 17.8; DB 10; Length 664;
Best Local Similarity 90.5%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 999ggacgatactgcggggg 21
|||||
Db 35 GGGGAGCATTTCGTCGTGG 55

RESULT 10
BI825162 677 bp mRNA linear EST 04-OCT-2001

LOCUS BI825162 603072034F1 NIH_MGC_119 Homo sapiens CDNA clone IMAGE:5163969 5',
DEFINITION mRNA sequence.

ACCESSION BI825162
VERSION BI825162.1 GI:15936712
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 677)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1407 row: c column: 10
High quality sequence stop: 677.
Location/Qualifiers

FEATURES

1..677

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5163969"

/clone_1ib="NIH_MGC_119"

/tissue_type="medulla"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: EcoRV (destroyed); RNA source normal medulla from

anonymous male age 27. Library is oligo-dT primed and

directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.3 kb, insert size range

0.9-3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 013. Note:

this is a NIH_MGC library."

BASE COUNT

148 a 159 c 230 g 140 t

ORIGIN

Query Match

Best Local Similarity

Matches 19; Conservative

OY 1 999ggacgatactgcggggg 21

|||||

Db 32 GGGGAGCATTTCGTCGTGG 52

RESULT 11

BI821280 686 bp mRNA linear EST 04-OCT-2001
LOCUS 603038109F1 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5178915 5',
DEFINITION mRNA sequence.

ACCESSION BI821280
VERSION BI821280.1 GI:15932830
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 686)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1446 row: b column: 04
High quality sequence stop: 686.
Location/Qualifiers

FEATURES

1..686

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5178915"

/clone_1ib="NIH_MGC_115"

/lab_host="DH10B"

/note="Organ: pooled brain, lung, testis; Vector:

pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA

source anonymous pool of 6 male brains, age range 23-27; 1

male lung, age 27; and 1 male testis, age 69. Library is

oligo-dT primed and directionally cloned (EcoRV site is

destroyed upon cloning). Average insert size 1.8 kb,

insert size range 1-3 kb. Library is normalized and

enriched for full-length clones and was constructed by C.

Gruber (Invitrogen). Research Genetics tracking code

021. Note: this is a NIH_MGC library."

BASE COUNT

163 a 161 c 224 g 138 t

ORIGIN

Query Match

Best Local Similarity

Matches 19; Conservative

OY 1 999ggacgatactgcggggg 21

|||||

Db 33 GGGGAGCATTTCGTCGTGG 53

RESULT 12

BI828300 692 bp mRNA linear EST 04-OCT-2001

LOCUS 603078196F1 NIH_MGC_119 Homo sapiens CDNA clone IMAGE:5170055 5',
DEFINITION mRNA sequence.

ACCESSION BI828300

VERSION BI828300.1 GI:15939850

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 692)

NIH-MGC http://mhc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-rt@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L14M1422 row: p column: 24
 High quality sequence stop: 692.
 Location/Qualifiers

FEATURES

source

1.692
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5170055"
 /clone_1id="NIH_MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."
 BASE COUNT 139 a 167 c 242 g 144 t
 ORIGIN

Query Match 80.9%; Score 17.8; DB 10; Length 692;
 Best Local Similarity 90.5%; Pred. No. 1.3e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gggggacgatactgcggggg 21
 ||||||| |||||||
 Db 60 GGGGACGATTTCTGCTGG 80

RESULT 13 694 bp mRNA linear EST 16-OCT-2001
 BI907688
 LOCUS 603065842F1 NIH_MGC_118 Homo sapiens CDNA clone IMAGE:5214682 5',
 DEFINITION mRNA sequence.
 ACCESSION BI907688
 VERSION BI907688.1 GI:16170530
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 694)
 NIH-MGC http://mgc.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-rt@mail.nih.gov

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L14M1539 row: d column: 11
 High quality sequence stop: 694.
 Location/Qualifiers

FEATURES

source

1.694
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5214682"
 /clone_1id="NIH_MGC_118"

/tissue_type="leukocyte"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."
 BASE COUNT 143 a 188 c 211 g 152 t
 ORIGIN

Query Match 80.9%; Score 17.8; DB 10; Length 694;
 Best Local Similarity 90.5%; Pred. No. 1.3e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gggggacgatactgcggggg 21
 ||||||| |||||||
 Db 34 GGGGACGATTTCTGCTGG 54

RESULT 14 702 bp mRNA linear EST 16-OCT-2001
 BI910445
 LOCUS 603067754F1 NIH_MGC_118 Homo sapiens CDNA clone IMAGE:5216797 5',
 DEFINITION mRNA sequence.
 ACCESSION BI910445
 VERSION BI910445.1 GI:16173834
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 702)
 NIH-MGC http://mgc.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-rt@mail.nih.gov

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L14M1544 row: l column: 14
 High quality sequence stop: 700.
 Location/Qualifiers

FEATURES

source

1.702
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5216797"
 /clone_1id="NIH_MGC_118"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."
 BASE COUNT 154 a 162 c 240 g 146 t
 ORIGIN

Query Match 80.9%; Score 17.8; DB 10; Length 702;

Best Local Similarity 90.5%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 99999999999999999999 21
1111111111111111111111
Db 39 GGGGAGCATTCGTCGTGG 59

RESULT 15

BE885598

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE885598 746 bp mRNA linear EST 20-OCT-2000
601508801F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3910177 5',
mRNA sequence.

BE885598
BE885598.1 GI:10334374
EST.

human.
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 746)

NIH-MGC <http://mgc.ncl.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>
Plate: LLM9725 row: b column: 02

High quality sequence stop: 642.

Location/Qualifiers

1..746

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3910177"

/clone_1ib="NIH_MGC_71"

/tissue_type="leiomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 2.1 kb.

BASE COUNT 177 a 175 c 243 g 151 t

ORIGIN

Query Match 80.9%; Score 17.8; DB 10; Length 746;
Best Local Similarity 90.5%; Pred. No. 1.3e+03;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 99999999999999999999 21
1111111111111111111111
Db 2 GGGGAGCATTCGTCGTGG 22

Search completed: August 10, 2002, 02:11:08
Job time: 13129 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:03:41 ; Search time 277.54 Seconds
(without alignments)
19.471 Million cell updates/sec

Title: US-09-672-126-9
Perfect score: 22
Sequence: 1 gggggacgatctcgggggg 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	16.2	73.6	1929	US-08-750-180-1	Sequence 1, Appli
2	16.2	73.6	2148	PCT-US93-01842-1	Sequence 1, Appli
3	16.2	73.6	5275	US-08-485-588-1	Sequence 1, Appli
4	16.2	73.6	5275	US-08-484-565-1	Sequence 1, Appli
5	16.2	73.6	5275	US-08-480-751-1	Sequence 1, Appli
6	16.2	73.6	5275	US-08-943-986-1	Sequence 1, Appli
7	16.2	73.6	5275	US-08-353-784-1	Sequence 1, Appli
8	16.2	73.6	5275	US-08-484-719B-1	Sequence 1, Appli
9	16.2	73.6	5275	US-08-484-159-1	Sequence 1, Appli
C 10	16.2	73.6	43804	US-09-171-461-1	Sequence 1, Appli
11	15.8	71.8	1621	US-09-013-881-14	Sequence 14, Appl
12	15.8	71.8	4403765	US-09-103-840A-2	Sequence 2, Appli
13	15.8	71.8	4411529	US-09-103-840A-1	Sequence 1, Appli
14	15.6	70.9	558	US-08-617-785-9	Sequence 12, Appl
C 15	15.6	70.9	573	US-08-290-665A-128	Sequence 128, App
C 16	15.6	70.9	573	PCT-US95-10398-128	Sequence 11, Appl
17	15.6	70.9	1912	US-08-868-435-11	Sequence 11, Appl
18	15.6	70.9	1912	US-08-744-231-11	Sequence 11, Appl
19	15.6	70.9	2327	US-08-868-435-1	Sequence 1, Appli
20	15.6	70.9	2327	US-08-744-231-1	Sequence 1, Appli
21	15.6	70.9	2745	US-08-617-785-11	Sequence 11, Appl
22	15.6	70.9	2766	US-08-617-785-13	Sequence 13, Appl
23	15.6	70.9	3809	US-08-485-588-3	Sequence 3, Appli
24	15.6	70.9	3809	US-08-484-565-3	Sequence 3, Appli
25	15.6	70.9	3809	US-08-480-751-3	Sequence 3, Appli
26	15.6	70.9	3809	US-08-943-986-3	Sequence 3, Appli
27	15.6	70.9	3809	US-08-353-784-3	Sequence 3, Appli

28	15.6	70.9	3809	3	US-08-484-719B-3	Sequence 3, Appli
29	15.6	70.9	3809	4	US-08-546-998-2	Sequence 2, Appli
30	15.6	70.9	3809	4	US-08-484-159-3	Sequence 3, Appli
31	15.6	70.9	4000	2	US-08-687-289A-2	Sequence 2, Appli
32	15.6	70.9	5006	1	US-08-485-588-2	Sequence 2, Appli
33	15.6	70.9	5006	1	US-08-484-565-2	Sequence 2, Appli
34	15.6	70.9	5006	2	US-08-480-751-2	Sequence 2, Appli
35	15.6	70.9	5006	2	US-08-943-986-2	Sequence 2, Appli
36	15.6	70.9	5006	3	US-08-353-784-2	Sequence 2, Appli
37	15.6	70.9	5006	3	US-08-484-719B-2	Sequence 2, Appli
38	15.6	70.9	5006	4	US-08-546-998-1	Sequence 1, Appli
39	15.6	70.9	5006	4	US-08-484-159-2	Sequence 2, Appli
C 40	15.2	69.1	1917	2	US-08-637-899-2	Sequence 2, Appli
C 41	14.8	67.3	23	3	US-08-772-512A-7	Sequence 7, Appli
C 42	14.8	67.3	1167	2	US-08-492-027A-5	Sequence 5, Appli
C 43	14.8	67.3	1329	3	US-08-360-758-1	Sequence 1, Appli
C 44	14.8	67.3	1389	1	US-08-458-023B-1	Sequence 1, Appli
C 45	14.8	67.3	1389	3	US-09-111-556A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-750-180-1/c
; Sequence 1, Application US/08750180
; Patent No. 6284880
; GENERAL INFORMATION:
; APPLICANT: COTTEN, MATTHEW
; APPLICANT: BAKER, ADAM
; APPLICANT: CHIOCCA, SUSANNA
; TITLE OF INVENTION: Method for Introducing Foreign Material into
; TITLE OF INVENTION: Higher Eukaryotic Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,180
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/01989
; FILING DATE: 26-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 18 825.0
; FILING DATE: 30-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 42 587.2
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RAZ E. FLESHNER
; REGISTRATION NUMBER: 34,331
; REFERENCE/DOCKET NUMBER: 0652.1580000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1929 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA

PCT-US93-01642-1

PCT-US93-01642-1

PCT-US93-01642-1

```

:
: ATTORNEY/AGENT INFORMATION:
:
:   NAME: Heber, Sheldon O.
:   REGISTRATION NUMBER: 38,179
:   REFERENCE/DOCKET NUMBER: 213/005
:
: TELECOMMUNICATION INFORMATION:
:
:   TELEPHONE: (213) 489-1600
:   TELEFAX: (213) 955-0440
:   TELEX: 67-3510
:
: INFORMATION FOR SEQ ID NO: 1:
:
:   SEQUENCE CHARACTERISTICS:
:
:     LENGTH: 5775 base pairs
:     TYPE: nucleic acid
:     STRANDEDNESS: single
:

```

TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
OTHER INFORMATION:
US-08-485-588-1

Query Match 73.6%; Score 16.2; DB 1; Length 5275;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ggggacgatctgcg9999g 22
||||| ||||| ||||| |||||
DB 605 GGGGACATTATCTCGGGGG 625

RESULT 4
US-08-484-565-1
; Sequence 1, Application US/08484565
; Patent No. 5763569
; GENERAL INFORMATION:
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,565
; FILING DATE: 7 June, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 9
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; FILING DATE: 21 August, 1992
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451
; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 213/006
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5275 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 515..3769
; OTHER INFORMATION:
; US-08-484-565-1

Query Match 73.6%; Score 16.2; DB 1; Length 5275;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ggggacgatctgcg9999g 22
||||| ||||| ||||| |||||
DB 605 GGGGACATTATCTCGGGGG 625

RESULT 5
US-08-480-751-1
; Sequence 1, Application US/08480751
; Patent No. 5858684
; GENERAL INFORMATION:
; APPLICANT: Edward F. Nemeth
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: Forrest H. Fuller
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,751
; FILING DATE: 7 June, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 9
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161

; FILING DATE: 21 August, 1992
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451
; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 213/004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5275 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 515..3769
; OTHER INFORMATION:
; US-08-480751-1

Query Match 73.6%; Score 16.2; DB 2; Length 5275;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 999gacgatctcgggggg 22
||||| ||||| ||||| |||||
Db 605 GGGGACATTATCTCGGGGG 625

RESULT 6
US-08-943-986-1
; Sequence 1, Application US/08943986
; Patent No. 5962314
; GENERAL INFORMATION:
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,986
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/484,565
; FILING DATE: 7-June-1995
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827

; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; FILING DATE: 21 August, 1992
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451
; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 213/006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5275 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 515..3769
; OTHER INFORMATION:
; US-08-943-986-1

Query Match 73.6%; Score 16.2; DB 2; Length 5275;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 999gacgatctcgggggg 22
||||| ||||| ||||| |||||
Db 605 GGGGACATTATCTCGGGGG 625

RESULT 7
US-08-353-784-1
; Sequence 1, Application US/08353784
; Patent No. 6011068
; GENERAL INFORMATION:
; APPLICANT: Edward F. Nemeth, Edward M.
; APPLICANT: Brown, Steven C. Hebert,
; APPLICANT: Bradford C. Van Wagenen, Manuel
; APPLICANT: F. Balandrin, Forrest H. Fuller,
; APPLICANT: Eric G. DelMar, and Scott T. Moe
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,784

;; FILING DATE: 9 December, 1994
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA: including application
;; PRIOR APPLICATION DATA: described below: 8
;; APPLICATION NUMBER: PCT/US/94/12117
;; FILING DATE: 21 October, 1994
;; APPLICATION NUMBER: U.S. 08/292,827
;; FILING DATE: 23 August, 1994
;; APPLICATION NUMBER: U.S. 08/141,248
;; FILING DATE: 22 October, 1993
;; APPLICATION NUMBER: U.S. 08/009,389
;; FILING DATE: 21 August, 1992
;; APPLICATION NUMBER: U.S. 07/834,044
;; FILING DATE: 11 February, 1992
;; APPLICATION NUMBER: U.S. 07/749,451
;; FILING DATE: 23 August, 1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Heber, Sheldon O.
;; REGISTRATION NUMBER: 38,179
;; REFERENCE/DOCKET NUMBER: 209/069
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 489-1600
;; TELEFAX: (213) 955-0440
;; TELEX: 67-3510
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5275 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: CDNA to mRNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 515..3769
;; OTHER INFORMATION:
US-08-353-784-1

Query Match 73.6%; Score 16.2; DB 3; Length 5275;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 ggggacgatctcggtgggg 22
||||| ||||| ||||| ||||| |||||
Db 605 GGGACATTATCTCGGGGG 625

RESULT 8
US-08-484-719B-1
; Sequence 1, Application US/08484719B
; Patent No. 6031003
; GENERAL INFORMATION:
; APPLICANT: Edward F. Nemeth, Edward M.
; APPLICANT: Brown, Steven C. Hebert,
; APPLICANT: Bradford C. Van Wagenen,
; APPLICANT: Manuel F. Baladrin,
; APPLICANT: Forrest H. Fuller, Eric G.
; APPLICANT: Delmar, Scott T. Moe
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California

;; COUNTRY: USA
;; ZIP: 90071
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: MS Word
;; SOFTWARE: FastSeq for Windows Version 3.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/484,719B
;; FILING DATE: 7 June, 1995
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/353,784
;; FILING DATE: 9 December, 1994
;; APPLICATION NUMBER: PCT/US/94/12117
;; FILING DATE: 21 October, 1994
;; APPLICATION NUMBER: U.S. 08/292,827
;; FILING DATE: 23 August, 1994
;; APPLICATION NUMBER: U.S. 08/141,248
;; FILING DATE: 22 October, 1993
;; APPLICATION NUMBER: U.S. 08/009,389
;; FILING DATE: 23 February, 1993
;; APPLICATION NUMBER: U.S. 08/017,127
;; FILING DATE: 12 February, 1993
;; APPLICATION NUMBER: U.S. 07/934,161
;; FILING DATE: 21 August, 1992
;; APPLICATION NUMBER: U.S. 07/834,044
;; FILING DATE: 11 February, 1992
;; APPLICATION NUMBER: U.S. 07/749,451
;; FILING DATE: 23 August, 1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Douglas C. Murdock
;; REGISTRATION NUMBER: 37,549
;; REFERENCE/DOCKET NUMBER: 213/007
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 489-1600
;; TELEFAX: (213) 955-0440
;; TELEX: 67-3510
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5275 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: CDNA to mRNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 515..3769
US-08-484-719B-1

Query Match 73.6%; Score 16.2; DB 3; Length 5275;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 ggggacgatctcggtgggg 22
||||| ||||| ||||| ||||| |||||
Db 605 GGGACATTATCTCGGGGG 625

RESULT 9
US-08-484-159-1
; Sequence 1, Application US/08484159
; Patent No. 6313146
; GENERAL INFORMATION:
; APPLICANT: Bradford C. Van Wagenen
; APPLICANT: Manuel F. Baladrin
; APPLICANT: Eric G. Del Mar
; APPLICANT: Edward F. Nemeth
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,159
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 214/101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515...3769
OTHER INFORMATION:
US-08-484-159-1

Query Match 73.6%; Score 16.2; DB 4; Length 5275;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ggggacgatctcgggggg 22
||||| ||||| ||||| |||||
Db 605 GGGGACATTTCTCGGGGG 625

RESULT 10
US-09-171-461-1/c
; Sequence 1, Application US/09171461
; Patent No. 6355016

GENERAL INFORMATION:
APPLICANT: Baker, Adam
APPLICANT: Cotten, Matthew
APPLICANT: Chioocca, Susanna
APPLICANT: Kurzbaue, Robert
APPLICANT: Schaffner, Gotthold
TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
FILE REFERENCE: 0652.180000
CURRENT APPLICATION NUMBER: US/09/171,461
CURRENT FILING DATE: 1999-01-12
EARLIER APPLICATION NUMBER: PCT/EP97/01944
EARLIER FILING DATE: 1997-04-18
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 43804
TYPE: DNA
ORGANISM: CELO Virus
FEATURE:
NAME/KEY: gene
LOCATION: (12193)..(15043)
OTHER INFORMATION: /gene: L1
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15080)
OTHER INFORMATION: /note= L2 region penton base splice acceptor site
FEATURE:
NAME/KEY: gene
LOCATION: (15110)..(17495)
OTHER INFORMATION: /gene: L2
FEATURE:
NAME/KEY: polyA_site
LOCATION: (17526)
FEATURE:
NAME/KEY: gene
LOCATION: (17559)..(21754)
OTHER INFORMATION: /gene: L3
FEATURE:
NAME/KEY: misc_feature
LOCATION: (18261)
OTHER INFORMATION: /gene: L3 /note= hexon splice acceptor site
FEATURE:
NAME/KEY: misc_feature
LOCATION: (21102)
OTHER INFORMATION: /gene: L3 /note= protease splice acceptor site
FEATURE:
NAME/KEY: misc_feature
LOCATION: (21123)
OTHER INFORMATION: /gene: L3 /note= protease splice acceptor site
FEATURE:
NAME/KEY: polyA_site
LOCATION: (21767)
FEATURE:
NAME/KEY: polyA_site
LOCATION: (21824)
FEATURE:
NAME/KEY: polyA_site
LOCATION: (21836)
FEATURE:
NAME/KEY: polyA_site
LOCATION: (21882)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23608)
OTHER INFORMATION: /note= 100K splice acceptor site
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23649)
OTHER INFORMATION: /note= 100K splice acceptor site
FEATURE:
NAME/KEY: gene
LOCATION: (23680)..(27886)
OTHER INFORMATION: /gene: L4

```
;
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: (27920)
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (28315)
; OTHER INFORMATION: /note= fibre splice acceptor site
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (28341)
; OTHER INFORMATION: / note= fibre splice acceptor site
; FEATURE:
; NAME/KEY: gene
; LOCATION: (28363)..(31768)
; OTHER INFORMATION: /gene: L5
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (30511)
; OTHER INFORMATION: /gene: L5 /note= fibre splice acceptor site
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: (31770)
US-09-171-461-1
```

```
Query Match 73.6%; Score 16.2; DB 4; Length 43804;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 ggggacgatctcgctggggg 22
||| ||||| ||||| |||||
DB 37946 GGAGCGATATTGTCGGGGG 37926
```

```
RESULT 11
US-09-013-881-14
; Sequence 14, Application US/09013881
; Patent No. 6132964
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,881
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0470 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
```

```
;
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1621 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: TESTNOT03
; CLONE: 2011230
US-09-013-881-14
```

```
Query Match 71.8%; Score 15.8; DB 3; Length 1621;
Best Local Similarity 89.5%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 ggggacgatctcgctgggg 20
||| ||||| ||||| |||||
DB 655 GGGGACGATATCGTGGCG 673
```

```
RESULT 12
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```

```
Query Match 71.8%; Score 15.8; DB 4; Length 4403765;
Best Local Similarity 89.5%; Pred. No. 41;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 ggggacgatctcgctgggg 20
||| ||||| ||||| |||||
DB 1151182 gcggacgatctcgctgggg 1151200
```

```
RESULT 13
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match          71.8%; Score 15.8; DB 4; Length 4411529;
Best Local Similarity 89.5%; Pred. No. 41;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ggggacgatctcgctggg 20
    ||||| |||||
Db 1151157 ggggacgatctcgctggg 1151175

RESULT 14
US-08-617-785-9
; Sequence 9, Application US/08617785E
; Patent No. 6228610
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Kropfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/08/617,785E
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: PCT/EP94/02991
; EARLIER FILING DATE: 1994-09-07
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(558)
US-08-617-785-9

Query Match          70.9%; Score 15.6; DB 4; Length 558;
Best Local Similarity 81.8%; Pred. No. 55;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gggggacgatctcgctgggg 22
    ||||| |||||
Db 138 gggggacgtaccctcggggg 159

RESULT 15
US-08-290-665A-128/c
; Sequence 128, Application US/08290665A
; Patent No. 582852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 573 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: T2
US-08-290-665A-128

Query Match          70.9%; Score 15.6; DB 2; Length 573;
Best Local Similarity 81.8%; Pred. No. 55;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gggggacgatctcgctgggg 22
    ||||| |||||
Db 317 GAGGGACGAGAACCTCGGGGG 296

Search completed: August 10, 2002, 03:06:03
Job time: 16029 sec.
```


THIS PAGE BLANK (0870)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:57:51 ; Search time 2778.35 Seconds
(without alignments)
165.704 Million cell updates/sec

Title: US-09-672-126-11

Perfect score: 22

Sequence: 1 gggggacgagctcggtggggg 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_un.*

28: em_vi.*

29: em_htg_hum.*

30: em_htg_inv.*

31: em_htg_inv.*

32: em_htg_inv.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

ALIGNMENTS

RESULT	1	AXI04798	Sequence 990 from Patent WO0122972	22 bp	DNA
LOCUS	AXI04798	Sequence 990 from Patent WO0122972			
DEFINITION	AXI04798	Sequence 990 from Patent WO0122972			
ACCESSION	AXI04798	Sequence 990 from Patent WO0122972			
VERSION	AXI04798.1	GI:13920995			
KEYWORDS		synthetic construct			
SOURCE		synthetic construct			
ORGANISM		artificial sequence			
REFERENCE		1 (bases 1 to 22)			
AUTHORS		Krieg, A.M., Schetter, C. and Vollmer, J.C.			
TITLE		Immunostimulatory nucleic acids			
JOURNAL		Patent: WO 0122972-A 990 05-APR-2001;			
		UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical			
		GmbH (DE)			
FEATURES		Location/Qualifiers			
		1..22			
		/organism="synthetic construct"			
		/db_xref="taxon:32630"			
BASE COUNT		2 a	4 c	14 g	2 t
ORIGIN					

```
Query Match      100.0%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggggacgagctcgtcgggggg 22
   |||||  |||||  |||||  |||||  |||||
Db 1 GGGGACGAGCTCGTCGGGGG 22

RESULT 2
AX105113
LOCUS AX105113 22 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 11 from Patent WO0122990.
ACCESSION AX105113
VERSION AX105113.1 GI:13921263
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
JOURNAL Patent: WO 0122990-A 11 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source Location/Qualifiers
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
misc_feature 1..2
/note="Backbone has phosphorothioate linkages."
misc_feature 3..16
/note="Backbone has phosphodiester linkages."
misc_feature 17..21
/note="Backbone has phosphorothioate linkages."
misc_feature 22
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 4 c 14 g 2 t
ORIGIN

Query Match      100.0%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggggacgagctcgtcgggggg 22
   |||||  |||||  |||||  |||||  |||||
Db 1 GGGGACGAGCTCGTCGGGGG 22

RESULT 3
AX104796
LOCUS AX104796 22 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 988 from Patent WO0122972.
ACCESSION AX104796
VERSION AX104796.1 GI:13920993
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 988 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source Location/Qualifiers
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"

Query Match      85.5%; Score 18.8; DB 6; Length 22;
Best Local Similarity 90.9%; Pred. No. 6.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgagctcgtcgggggg 22
   |||||  |||||  |||||  |||||  |||||
Db 1 GGGGACGATATCGTCGGGGG 22

RESULT 4
AX104797
LOCUS AX104797 22 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 989 from Patent WO0122972.
ACCESSION AX104797
VERSION AX104797.1 GI:13920994
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 989 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES
source Location/Qualifiers
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"

Query Match      85.5%; Score 18.8; DB 6; Length 22;
Best Local Similarity 90.9%; Pred. No. 6.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgagctcgtcgggggg 22
   |||||  |||||  |||||  |||||  |||||
Db 1 GGGGACGAGCTCGTCGGGGG 22

RESULT 5
AX105111
LOCUS AX105111 22 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 9 from Patent WO0122990.
ACCESSION AX105111
VERSION AX105111.1 GI:13921261
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
JOURNAL Patent: WO 0122990-A 9 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source Location/Qualifiers
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
misc_feature 1..2
/note="Backbone has phosphorothioate linkages."
misc_feature 3..16
/note="Backbone has phosphodiester linkages."
misc_feature 17..21
/note="Backbone has phosphorothioate linkages."
misc_feature 22
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 4 c 14 g 2 t
ORIGIN
```


misc_feature 22 /note="Backbone has phosphorothioate linkages."
BASE COUNT 3 a 3 c 13 g 3 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 6; Length 22;
Best Local Similarity 90.9%; Pred. No. 6.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgagctgcgtcgggggg 22
||||| ||||| ||||| |||||
Db 1 GGGGACGATATCGTCGGGGG 22

RESULT 6
AX105112 AX105112 22 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 10 from Patent WO0122990.
DEFINITION AX105112
ACCESSION AX105112
VERSION AX105112.1 GI:13921262
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced interferon

JOURNAL
Patent: WO 0122990-A 10 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)

FEATURES
source
1..22 Location/Qualifiers
/organism="synthetic construct."
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
1..2
misc_feature
/note="Backbone has phosphorothioate linkages."
3..16
misc_feature
/note="Backbone has phosphodiester linkages."
17..21
misc_feature
/note="Backbone has phosphorothioate linkages."
22
misc_feature
/note="Backbone has phosphodiester linkages."
2 a 4 c 14 g 2 t
BASE COUNT 2 a 4 c 14 g 2 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 6; Length 22;
Best Local Similarity 90.9%; Pred. No. 6.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgagctgcgtcgggggg 22
||||| ||||| ||||| |||||
Db 1 GGGGACGAGCTGCTCGGGGG 22

RESULT 7
AB032366
LOCUS Mus musculus Ehm2 mRNA, complete cds.
DEFINITION AB032366 3451 bp mRNA linear ROD 23-MAY-2000
ACCESSION AB032366
VERSION AB032366.1 GI:8051691
KEYWORDS
SOURCE Mus musculus
ORGANISM Mus musculus cDNA to mRNA.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Hashimoto,Y., Shindo-Okada,N., Tani,M., Takeuchi,K., Toma,H. and Yokota,J.

TITLE Identification of genes differentially expressed in association with metastatic potential of K-1735 murine melanoma by messenger RNA differential display
JOURNAL Cancer Res. 56 (22), 5266-5271 (1996)
MEDLINE 97069887
REFERENCE
AUTHORS Shimizu,K., Nagamachi,Y., Tani,M., Kimura,K., Shirolshi,T., Wakana,S. and Yokota,J.
TITLE Molecular cloning of a novel NF2/ERM/4.1 superfamily gene, ehm2, that is expressed in high-metastatic K1735 murine melanoma cells
JOURNAL Genomics 65 (2), 113-120 (2000)
MEDLINE 20247250
REFERENCE 3 (bases 1 to 3451)
AUTHORS Yokota,J., Shimizu,K. and Nagamachi,Y.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-1999) Jun Yokota, National Cancer Center Research Institute, Biology Division; Tsukiji 5-chome 1-1, Chuo-ku, Tokyo 104-0045, Japan (E-mail:jyokota@ncc.ncc.go.jp, Tel:81-3-3547-5272, Fax:81-3-3542-0807)

FEATURES
source
1..3451 Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="4"
499..2082
/gene="Ehm2"
499..2082
/gene="Ehm2"
/codon_start=1
/product="Ehm2"
/protein_id="BA096078.1"
/db_xref="GI:8051682"
/translation="MLRFLRRTFGRSMORYARGAAGRGAGLDERDGGPRGGPAAASSSVLPAPGGVPPAGGGLLTGGAHVHISAGAAKATLYCRVFLLDGTEVSDLPKXAKQDLFDQIVYHLDLVETDFGLQFLDSQVTHWDHAKPIKKOMKVGAYALHF RVKYSSEPNNREEFRYLFVQLRHLILSGLKACPETAVELAAALCKVQELGECGL PEHTLSEVSEFRFPNQTEAMEFDIFQRMKEYGKSPQAELSYLNKAKWLEMYGDM HVVRGDGCEYSLGTLPTGILIFEGAKGILGFMPKTKMDKFKSKLTLLVVEDDQGR REQHTFVPLDSARTCKHLWKCAVEHFAFFRLRTPSNSKASRSDFLRSGSRFSGR TEYQATHGSLRRTSTFERKPSKRYPSRRHSTFKASNPITIAAOLCSKANPEVHNYQPYHPDVHPSQPRWRHPSPNYSICKQKPCQDDRPWKTSASGDSHFDDYQDQNRNLGGASVTYRDKLMTAL"

BASE COUNT 764 a 958 c 967 g 762 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 10; Length 3451;
Best Local Similarity 90.9%; Pred. NO. 1.8e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgagctgcgtcgggggg 22
||||| ||||| ||||| |||||
Db 585 GGGGACGAGCGCGACGGGGG 606

RESULT 8
AE005862
LOCUS Caulobacter crescentus section 188 of 359 of the complete genome.
DEFINITION AE005862 AE005673
ACCESSION AE005862.1 GI:13423326
VERSION
KEYWORDS
SOURCE
ORGANISM Caulobacter crescentus.
Bacteria; Proteobacteria; alpha subdivision; Caulobacter group; Caulobacter.
REFERENCE 1 (bases 1 to 10016)
AUTHORS Nierman,W.C., Feldblyum,T.V., Laub,M.T., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R., Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Deboy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L., Haft,D.H., Kolonay,J.F., Smit,J., Craven,M., Khouri,H., Shetty,J., Berry,K., Utterback,T., Tran,K., Wolf,A., Vamathevan,J.,

TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS

Ermolaeva,M., White,O., Salzberg,S.L., Venter,J.C., Shapiro,L. and Fraser,C.M.
 Complete genome sequence of *Caulobacter crescentus*
 Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)
 21173698

2 (bases 1 to 10016)
 Nierman,W.C., Feldblyum,T.V., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R., Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Laub,M.T., DeBoy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L., Haft,D.H., Kolonay,J.F., Smit,J., Craven,M., Khouri,H., Shetty,J., Berry,K., Utterback,T., Iran,K., Wolf,A., Vamathevan,J., Ermolaeva,M., White,O., Salzberg,S.L., Shapiro,L., Venter,J.C. and Fraser,C.M.
 Direct Submission
 Submitted (31-JAN-2001) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
FEATURES
 source

1. .10016
 /organism="Caulobacter crescentus"
 /db_xref="taxon:69394"
 complement(86..895)
 /gene="CCI180"
 complement(86..895)
 /gene="CCI180"
 /note="similar to GB:M88701 SP:P30139 PID:409790
 PID:414235 GB:U00096; identified by sequence similarity;
 putative"
 /codon_start=1
 /transl_table=11
 /product="thig protein"
 /protein_id="AAK23855.1"
 /db_xref="GI:13423327"
 /translation="MAHVTPTSDTETVAGTRFSRLIVGTGKDYKDYATNAA
 ARAAGAEIVTVAVRNVLTDPSPQLLDYVVKPTFLPNTAGTCGEDAVRRLRA
 REAGNDVLKLEISDPKTLFPDMEETLRSLKLVADGQVNVYCGDDPVYARLKEA
 GAVIMPILGAGISGLGIGNRVNLRIIENAKVPVLVDAGVGTASDAIGAHELCDGA
 LMNTATAEADKPIPMAKAMKHAIVIAGREAYLAGRMQKRLYADPPSSPLAGLI"
 complement(1040..1240)
 /gene="CCI181"
 complement(1040..1240)
 /gene="CCI181"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAK23856.1"
 /db_xref="GI:13423328"
 /translation="MRLLNGEERDVAGVVTIADLVTLGLDARKVAVERNLEIAPRS
 SYATALADGRIEIVTFITGG"
 1319..1759
 /gene="CCI182"
 1319..1759
 /gene="CCI182"
 /note="identified by match to PFAM protein family HMM
 PF01220"
 /codon_start=1
 /transl_table=11
 /product="3-dehydroquinase dehydratase, type II"
 /protein_id="AAK23857.1"
 /db_xref="GI:13423329"
 /translation="MVKPTHLVLSGPNLNLGTRPEIYKGTLDQVTRCFEARSGRS
 VSVVFRQSNHEGVLIDWVHEARSAALVNPAGHTSIALLDLTLSPVTECHH
 SNPAAREFRHHYVLSAATGIYSGFAASYLEATAEAFGLIRA"
 1820..2332
 /gene="CCI183"
 1820..2332
 /gene="CCI183"
 /note="identified by match to TIGR protein family HMM
 TIGR00531"
 /codon_start=1
 /transl_table=11
 /product="acetyl-CoA carboxylase, biotin carboxyl carrier

```
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK23863.1"
/db_xref="GI:13423335"
/translation="MLKCAATGARLDCEGAGRYGPPASRLSRLSKALQIQVRT"
5602..8415
/gene="CC1889"
5602..8415
/gene="CC1889"
5602..8415
/note="similar to GP:3402271; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ribonucleotide reductase-related protein"
/protein_id="AAK23864.1"
/db_xref="GI:13423336"
/translation="MRPYAHGRVESHVRLDVLVTHAGISDLETMPRPSAKAAGLVPAM
EMRLERADSVQVVSAPSSDSADVEAWLDMGTEPEKPIGGGPPARYAARVAQILGLS
RGLFGDADADAREALLTMTGTATPTSGOTGLRLLPDITEREFOSQARDILGQRR
ARLAAQAAALNTALDQVADAVRSCDSRACSDVRKNPALAARAKARELGADQOT
ILDAIALADQPRAPLTALANDDELPLVAANTROQVAAGDEAATLAAQVWETSALTIA
LSPDDAELSAGAAIGAADSTFDAGFDIERFTFVCHLWATALELDGERPARLS
LAGVDCLLSQYQATSTIEGRGAAALMALAGALSASAEAAATLGQDPADFDDROG
VLKTLAERRVRAALRSALATEAAALATAHALAKHGLRSAGLVGPADLEATLRLG
GSPVGAAGMAPVSVOTADGFLMPSFAAFEALESSEGVLDAAARRHALGHSIAES
PAIDHMLQAGFTAHETEKAENALREHOGVRAAPAPAIACAGFLRDVLGASAEVAR
RDFDTLAFDSAEIAAERHAGLGTLDGCEALNDPLREAFRSVPSNDRMLAM
AAVSFECLPTAVTSIPHDRAADAVRAQAAARAGVRLRLQRAETPYDFKLDLDE
EYDEAPAPAPREFLTVRVERIVERDRSRRLPRRKGIQKAAVGGHKVLIHTG
EYDEGELCELFIDMKGAAPRSLMNNFAIIVSLGLOHVPDFVAFVYTKPEPAG
PVTGNDTSKATSLTIDYIFRELGVSLGRDLDNGDQGFNADGLGAKKRAEADLLD
DEADPVASRFLSKGSEGPATPDNLIVFASFGHRRVEGADRPGADGEMCPACGDLISLR
RGMIVCDTCGASRPEPVAST"
8487..8984
/gene="CC1890"
8487..8984

Query Match      85.5%; Score 18.8; DB 1; Length 10016;
Best Local Similarity 90.9%; Pred. No. 1.4e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgagctcgcggggg 22
|||||
Db 4119 GGGGACGAGCTGTCGCGGG 4140

RESULT 9
SC9H11/c
LOCUS SC9H11 26500 bp DNA linear BCT 22-MAY-2000
DEFINITION Streptomyces coelicolor cosmid 9H11.
ACCESSION AL356592
VERSION AL356592.1 GI:8052359
KEYWORDS Arac-family transcriptional regulator; bldA codon; dioxxygenase;
DNA-binding protein; hydrolase; membrane efflux protein; membrane
protein; NAD-dependent dehydratase; narB; nitrate reductase;
oxidoreductase; TetR-family transcriptional regulator;
transmembrane transport protein.
SOURCE Streptomyces coelicolor A3(2).
ORGANISM Streptomyces coelicolor A3(2)
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Streptomyces.
REFERENCE 1 (bases 1 to 26500)
AUTHORS Redenbach,M., Kleser,H.M., Denapaita,D., Eichner,A., Cullum,J.,
Kinashi,H. and Hopwood,D.A.
TITLE A set of ordered cosmids and a detailed genetic and physical map
for the 8 Mb Streptomyces coelicolor A3(2) chromosome
JOURNAL Mol. Microbiol. 21 (1), 77-96 (1996)
MEDLINE 97000351
REFERENCE 2 (bases 1 to 26500)
AUTHORS Seeger,K.J. and Harris,D.
JOURNAL Unpublished
```

```
REFERENCE 3 (bases 1 to 26500)
AUTHORS Thomson,N.R., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
TITLE Direct Submission
JOURNAL Submitted (19-MAY-2000) Streptomyces coelicolor sequencing project,
Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
Colney, Norwich, Norfolk NR4 7UH, UK
COMMENT Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded
by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are
available on the World Wide Web.
(URL; http://www.sanger.ac.uk/Projects/S_coelicolor/)
CDS are numbered using the following system eg SC7B7.01c. SC (S.
coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for
CDS which show significant similarity to other CDS in the database.
The position of possible ribosome binding site sequences are given
where these have been used to deduce the initiation codon.
Gene prediction is based on positional base preference in codons
using a specially developed Hidden Markov Model (Krogh et al.,
Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot
program of Bibb et al., Gene 30:157-66(1984) as implemented at
http://www.nhn.gu.jp/
jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the
correct initiation codon. Where possible we choose an initiation
codon (atg, gtg, ttg or (att)) which is preceded by an upstream
ribosome binding site sequence (optimally 5-13bp before the
initiation codon). If this cannot be identified we choose the most
upstream initiation codon.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions.
Cosmid 9H11 lies between and overlaps cosmids 4G10 and 10G8 on the
AseI-A genomic restriction fragment.
FEATURES
source
1..26500
/organism="Streptomyces coelicolor A3(2)"
/strain="A3(2)"
/db_xref="taxon:100226"
/clone="cosmid 9H11"
misc_feature
1..106
/gene="SC9H11.01"
/note="nominal overlap with cosmid St4G10 between bases
31231..31336."
CDS
<1..607
/gene="SC9H11.01"
/note="SC9H11.01, unknown, partial CDS, len: >201 aa."
/codon_start=2
/transl_table=11
/product="hypothetical protein SC9H11.01."
/protein_id="CAB92190.1"
/db_xref="GI:8052360"
/translation="ILTTAERLFAEHGVYAVSNROVSEAGGNNAAVGYHFGTKTDL
VRAIAQRHSEVEELARQLALGSDPLRDWDCLVRFQPDHLAALGSPWYARFCA
QVMTDPAQLQIMTESRASVSLRAIIVGNRCMPALPDVEAERAGDMARHLIVTAEE
RERAAENRPTPRASQWQADAGLDVAIVGMVLAPVTPRGGG"
1..607
/gene="SC9H11.01"
774..808
/note="SC9H11 repeat unit 1 (RU1). Repeated three times on
this cosmid (iterated at positions 965..999 and
1156..1190) with the consensus:
GCTG(C/T)GAGGGCGCGGGCGGCGCGGTCCA."
809..815
/note="SC9H11 repeat unit 2 (RU2). Repeated three times
(alternate positions 1000..1006 and 1192..1198) with the
```


source

1. 135756
 /organism="Oryza sativa"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /chromosome="8"
 /clone="P0020B10"

BASE COUNT 38394 a 29299 c 29006 g 38857 t 200 others
 ORIGIN

Query Match 85.5%; Score 18.8; DB 2; Length 135756;
 Best Local Similarity 90.9%; Pred. No. 7.3e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 gggggacgagctcgtcggggg 22
 |||||
 Db 17058 GGGGACGAGCTCGTGGCTGG 17079

RESULT 11

AP003946 138906 bp DNA linear HTG 26-JUL-2001
 LOCUS Oryza sativa chromosome 6 clone OJ1147_D11, *** SEQUENCING IN
 DEFINITION PROGRESS ***, in ordered pieces.

ACCESSION AP003946
 VERSION AP003946.1 GI:15021916
 KEYWORDS HTG: HTGS_PHASE2

SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:OJ1147_D11.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 138906)
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
 TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, BAC clone:OJ1147_D11

JOURNAL Published Only in Database (2001) In press

REFERENCE 2 (bases 1 to 138906)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Direct Submission
 JOURNAL Submitted (25-JUL-2001) Takuji Sasaki, National Institute of
 Agrobiological Resources, Rice Genome Research Program; Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)

COMMENT The nucleotide sequence of this BAC clone was generated by
 combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES

Location/Qualifiers

1..138906
 /organism="Oryza sativa"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /chromosome="6"
 /clone="OJ1147_D11"

BASE COUNT 39967 a 29996 c 29997 g 38791 t 155 others
 ORIGIN

Query Match 85.5%; Score 18.8; DB 2; Length 138906;
 Best Local Similarity 90.9%; Pred. No. 7.2e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 gggggacgagctcgtcggggg 22

Db 68401 GGGGACGAGCTCGTGGCTGG 68422
 |||||

RESULT 12

AP002865/c 139399 bp DNA linear PLN 26-JAN-2001
 LOCUS Oryza sativa genomic DNA, chromosome 1, PAC clone:P0034C11.
 DEFINITION AP002865
 ACCESSION AP002865
 VERSION AP002865.1 GI:10179050

SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:P0034C11.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 139399)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone:P0034C11

JOURNAL Published Only in Database (2000) In press

REFERENCE 2 (bases 1 to 139399)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Direct Submission

JOURNAL Submitted (13-SEP-2000) Takuji Sasaki, National Institute of
 Agrobiological Resources, Rice Genome Research Program; Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)

COMMENT Genes were predicted from the integrated results of the following:
 GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
 (October 1998 version). The genomic sequence was searched against
 NCBI NonRedundant Protein database, nr
 (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
 RGP. Protein homologies of the coding regions were searched against
 NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
 the identified cDNA sequences using BLASTN 2.0 with the
 corresponding DBJ accession no. and RGP clone ID.

A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with EST homology (covering
 almost the entire length of partial sequence) is classified as an
 'unknown' protein. A gene predicted with a gene prediction program
 is classified as a 'hypothetical' protein.
 The orientation of the sequence is from T7 to SP6 of the PAC clone.
 This sequence of P0034C11 clone has an overlap with P0434D08 (DBJ:
 AP001278) clone at the position 123,788 to 139,399 of 3' end. The
 sequence of this clone ends at the position 15,612 of P0434D08.
 Detailed information on overlap and assembly quality together with
 annotation of this entry is available at
 http://rgp.dna.affrc.go.jp/GenomeSeq.html.

FEATURES

source

1..139399
 /organism="Oryza sativa"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /chromosome="1"
 /clone="P0034C11"
 join(226..467,1083..1278,1372..1476,1688..1867,1979..2308,
 2397..2510,2660..3400)
 /gene="P0034C11.1"
 join(226..467,1083..1278,1372..1476,1688..1867,1979..2308,
 2397..2510,2660..3400)
 /gene="P0034C11.1"
 /note="contains EST C26525(C12525)"
 /codon_start=1
 /product="putative WRKY DNA binding protein"

/protein_id="BAB18313.1"
 /db_xref="GI:11320830"
 /translation="MLTSIFLPTCTPASALVPPETESITVVVVDIHRVRRSGRVRC
 TVHLFVQMYTKPLSSSYVWASDSATVDGMMVDENRSTMEADSGGGGGRRRVSVEV
 DFFSDEKKNMKKRSYGGVAAEADDAKPAAGLAIKKEDLTINLLPAGNNARSDRSM

gene

CDS

```
VVDDAASRPDHEKSSNELAAMQAEIGRMNEENORLRGMLTOVTTSYQALOMHLV
ALMOORPQMOPPTOPEPPHQDCKAEGAVVPROFLDLGPPSGAGGEAAEPESSNST
EAGSPRSSSTGNKDERGSDPAPSTAAWLPGRAMAPQMGAAAGKSHDQAOADA
NMKRAVYRSEASEPIIADGCOMRYKQGMAGNCPPRAYRYCTMATGCPVKRQVQR
CAEDRSILVTTEGTHHLPPLPAAMASTTAAASMLLSGMPADGAAGLMSNFL
ARTVLPCSSSMATISASAPFPVTLDLTHAPGAPNAVPLNAARCAPAPQFOVPLPG
GMAPAFVPOVLVYNOSEFSGLOMSDSEAAAAAQAQPPPPGIGOLGPLSD
TVSARAAITADPNFTVALAAITSIIGQHAAGNANNNTNTNTSNTNTNTSSNN
TVSNNTNSETQ"
join(4693. .5020,5085. .5211,8132. .8216)
/genes="P0034C11.2"
CDS
join(4693. .5020,5085. .5211,8132. .8216)
/genes="P0034C11.2"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB18314.1"
/db_xref="GI:11320831"
/translation="MWGAAYRMREKSWGARRCGAHPLNTRKPSQSALRPASRWYGT
ACVTVRYLTRSPYGRGVDFDRSSSPRAAGPRVVRERROKSGSLIVFWRYLHGELRA
SYVAGAGEVYASAGNAGGAGRERGFQVYKRLCTRIYKQGEVFRHAWHLSVN
HIVHARIPHHKIDIAFSG"
complement(9603. .10320)
/notes="3' LTR"
complement(10317. .13979)
/genes="P0034C11.3"
CDS
complement(10317. .13979)
/genes="P0034C11.3"
/codon_start=1
/product="putative polyprotein"
/protein_id="BAB18315.1"
/db_xref="GI:11320832"
/translation="WYVTRNGRTTGEBSNGEERTDGVHPNSDSNGNGPPPLPENPTLA
QYMAHOTMMTAMMOQHOQMYORMQQAEOHQOQPPPOSKLPEFLVRPPT
FESTTNPKANDLHAIKKLLNQNDQEKVAFATHQLOGPASWNDHNTTRPPT
EYVTAEFNRKFAOVLGCVAAQKREFRALHOGNRTVEYLHEDFLARVAPEDVRT
DAKQKFLAGLDDELTLNOLISGVADFERLVDAKAIROEDOHNMKDRKKAQFESNQ
ASHQRPRIPOQGEPTTMIVRQHPFNSSPHQASQSNHRGQGNRSATPRPMA
PAQSAFQVAKETGAKGSCFNGELGHFADKCPKPRRAGPRVQVARNHASTEEAQ
AAPEVLVTFPVNSIPATVDFSGAMHSFYKFKVGMHGLIREELSTPMRVTGNSS
TSQVSPSTVIEQSPFLANLILLESKDLDVILGMDLWTFKGVIDCANCTVTLINE
KGVTVYKSLVSPKQVSLNQIEVIVPVTTEKNSRKLKEIPVCEYEPVEPDELTTM
PKRIERFDLAPGATPIYKRPYRMAANELAELVKKQVDEOLQKGYIRPSTSPGAPV
IFVEKKDKTKMCDVRYALNEVTIKNKYLPRIIDLFLQGLAKGVSKIDLSGVHQL
RIRERDIPKTAFTQYGLYECTVMSFGLTNAPAFNMNKNVMEFLDKFVVVFIDDI
LIYSKLEEEHQLRLVLEKLEHQLYAKSKCDFWLSKVFLGHIITAGQVAVDSSN
LVSFTTTPKPTVPSYIRSEFLAGYRRIENFSRIARPMTQLLKDKFKWTAECDK
SFEELKKLVSAVPLILPDQMKDFQVYCDASHGLGCVLMQEGRRVAYASQLRPHEG
NYPTHDLAVVVAHLKIWRHYLIGNRCEVYTDHKSLEYIFTPDNLNRQWRLELIK
DYDMSIHYHPGKANVADALSKSYCNALCTEDMCEELQOELERLNVGIVBHFVAAL
EARPTLVQFRAQVNDPEIVEIKKNRVKARDFLEDEHRTIWMGERLVCVDPDKELK
DLILTEARQTSIHPGSTKMTQDLKRFVWVSMRREIAEFVALGCDVQCORVKAHQRP
VGLFQPLQIPKWKNEEIGMDFITGLPRTSSGHDSTWVYVVDRLTKVAHFIPVHTVTGK
KLAELCLARIEFDLLR"
complement(14231. .15091)
/genes="P0034C11.4"
CDS
complement(14231. .15091)
/genes="P0034C11.4"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB18316.1"
/db_xref="GI:11320833"
/translation="MTLVGVGNAPPEYTVPLPSGELPHRVHLEVHGVTGLANVVV
EASGCTADHACQEAAYLMARLERHNHIFHTAYRFPFRASGDVSTFRPTVGENN
TTFGHMCVAMVMDHSDHLKASKALNDGLVRIIALKDEIARKKENAQLGLPAP
GGVIRITPTPKSTAPYIOLAKPNPPPPAPAPAPAPVPPVSAFALSAPASA
VRGPASNGWGLPATPGSRDHRSSSSSGTEPGSGETYLPDSRSRSEGGPGKQEDAFD
FTDRRSCSEH"
join(16858. .16860,17327. .17683,17830. .17924,18192. .18468)
/genes="P0034C11.5"
CDS
join(16858. .16860,17327. .17683,17830. .17924,18192. .18468)
/genes="P0034C11.5"
/notes="hypothetical protein"
/codon_start=1
```

```
/protein_id="BAB18317.1"
/db_xref="GI:11320834"
/translation="NRRTGGRSNVDWDHAGGPPPVHIGDPDRPGADRTVVRMRPGSAR
LGLAQRAPAGHGKGGCARLPATATGCRWRKGGGARFGRRLDRLPARLGGRRGRIEE
RCLTSGRRGGRRRTATCEGNRAPATLDLGGGADWVLDLHANPTAATARAGERERG
SGRRRGREGDVDDVARARARAGATAVAGAAASGNFLGEDPTGGPHLSYTPRRRGRG
LSWAVASGRPSKAGRRGNGPAH"
complement(19452. .20198)
/genes="P0034C11.6"
CDS
complement(19452. .20198)
/genes="P0034C11.6"
/notes="hypothetical protein
probably inactive due to stop codon in CDS
similar to Oryza sativa chromosome 6, P0541H01.5"
/codon_start=1
/pseudo
complement(20254. .20971)
/notes="5' LTR"
join(24537. .24923,25850. .26024,26319. .26333)
/genes="P0034C11.7"
CDS
join(24537. .24923,25850. .26024,26319. .26333)
/genes="P0034C11.7"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB18318.1"
/db_xref="GI:11320835"
/translation="MKQRSDRNGKKTNEPRRCRHANDVACVDRDAAAHAHVGIFFAR
TPADPRSDAATLVRSMDNATEADKATVTDGDDGGGGGAGGAGGIARGSNGGG
EARGGAGHTARGDDDTGGGGGSHRSGRRREARGRRRRRGGEARGGPQRGKNDM"
complement(join(32273. .35615,35995. .36088))
/genes="P0034C11.8"
CDS
complement(join(32273. .35615,35995. .36088))
/genes="P0034C11.8"
/notes="hypothetical protein
probably inactive due to frameshift in CDS
similar to Oryza sativa chromosome 6, P0675A05.27"
/codon_start=1
/pseudo
complement(join(38816. .39184,39898. .40020))
/genes="P0034C11.9"
CDS
complement(join(38816. .39184,39898. .40020))
/genes="P0034C11.9"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB18319.1"
/db_xref="GI:11320836"
/translation="WVPFISLYOHCRRHWHVKGATAVELDLGSARYLGVKTVNG
LTIAADLTRKDEAGGRNATPEARGHICTIRHQACRRPPPSRPLPRAAATTKHCL
TVPLPLDHPDSFRLLRANPSAPPQLSTTASSYHRHRLPLPPLAESRERERE
REK"
join(41070. .41117,41126. .41249,41387. .41577,41605. .41994)
/genes="P0034C11.10"
CDS
join(41070. .41117,41126. .41249,41387. .41577,41605. .41994)
/genes="P0034C11.10"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB18320.1"
/db_xref="GI:11320837"
/translation="MDVALSGTVSSCHPGTRAVRVGCRGEVDPPTRRHNTESGDVW
SSSESMAGRRNPRGTADVAVAPLLVLVLDGKVQVQRRVYGLGGGVASEGKGGCGNH
```

```
Query Match 85.5%; Score 18.8; DB 8; Length 139399;
Best Local Similarity 90.9%; Pred. No. 7.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 99gggacgagctcgctgggggg 22
|||||
```

```
Db 118187 GGGGACGAGCTCGCGTGG 118166
|||||
```

```
RESULT 13
AC079356
LOCUS AC079356 174289 bp DNA linear HTG 29-AUG-2000
```

DEFINITION Oryza sativa chromosome 5 clone P0016H04, *** SEQUENCING IN
 PROGRESS ***, 9 ordered pieces.
 ACCESSION AC079356
 VERSION AC079356.1 GI:9937663
 KEYWORDS HTG; HTGS_PHASE2.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 (bases 1 to 174289)
 Hsing,Y.-I.C., Chow,T.-Y., Chen,C.-S., Wu,H.-P., Chao,Y.-T. and
 Liu,S.-M.
 TITLE Oryza sativa PAC P0016H04 genomics sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 174289)
 AUTHORS Hsing,Y.-I.C. and Chow,T.-Y.
 TITLE Direct Submission
 JOURNAL Submitted (29-AUG-2000) Institute of Botany, Academia Sinica, 128,
 Section 2, Yen-chu-Yuan Road, Nankang, Taipei 11529, Taiwan
 COMMENT * NOTE: This is a 'working draft' sequence. It currently
 * consists of 9 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 44682 contig of 44682 bp in length
 * 44683 93977 contig of unknown length
 * 93978 gap of unknown length
 * 93978 94928 contig of 951 bp in length
 * 94929 gap of unknown length
 * 94929 99763 contig of 4835 bp in length
 * 99764 105975 contig of unknown length
 * 105976 gap of unknown length
 * 105976 107680 contig of 1705 bp in length
 * 107681 gap of unknown length
 * 119112 149716 contig of unknown length
 * 149717 174289 contig of 30605 bp in length
 * 174289 gap of unknown length
 * 174289 contig of 24573 bp in length.
 FEATURES
 source Location/Qualifiers
 1..174289
 /organism="Oryza sativa"
 /db_xref="taxon:4530"
 /chromosome="5"
 /clone="P0016H04"
 BASE COUNT 51295 a 36890 c 36017 g 50285 t 2 others
 ORIGIN
 Query Match 85.5%; Score 18.8; DB 2; Length 174289;
 Best Local Similarity 90.9%; Pred. No. 6.8e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 gggggagcagctgcgtcgggggg 22
 |||||
 Db 160740 GGGGGAGCAGCTGCTGGTGG 160761
 RESULT 14
 AC007789/c
 LOCUS
 DEFINITION 182756 bp DNA linear PLN 03-DEC-1999
 Oryza sativa BAC OSJNBA0049B20 genomic sequence, complete sequence.
 ACCESSION AC007789
 VERSION AC007789.1 GI:5042437
 KEYWORDS HTG.
 SOURCE Oryza sativa.

ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 (bases 1 to 182756)
 Buehl,R., Benito,M.-I., Lin,X., Mason,T.M., Umayam,L., Shea,T.P.,
 Fujii,C.Y., Shen,M. and Fraser,C.M.
 TITLE Oryza sativa BAC OSJNBA0049B20 genomic sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 182756)
 AUTHORS Benito,M.-I.
 TITLE Direct Submission
 JOURNAL Submitted (11-JUN-1999) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 REFERENCE 3 (bases 1 to 182756)
 AUTHORS Benito,M.-I.
 TITLE Direct Submission
 JOURNAL Submitted (15-JUN-1999) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 REFERENCE 4 (bases 1 to 182756)
 AUTHORS Benito,M.-I.
 TITLE Direct Submission
 JOURNAL Submitted (18-JUN-1999) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 REFERENCE 5 (bases 1 to 182756)
 AUTHORS Benito,M.-I.
 TITLE Direct Submission
 JOURNAL Submitted (03-DEC-1999) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 COMMENT Address all correspondence to:
 Robin Buell or Maria-Ines Benito
 The Institute for Genomic Research
 9712 Medical Center Dr.
 Rockville, MD 20850, USA
 e-mail:rbuell@tigr.org or mbenito@tigr.org
 BAC clone OSJNBA0049B20 is from Oryza sativa.
 The orientation of the sequence is from SP6 to T7 end of the BAC
 clone.
 Genes were identified by a combination of three methods: Gene
 prediction programs including GENE (available by anonymous ftp
 from arthur.epm.ornl.gov), GeneFinder (Phil Green, University of
 Washington), Genscan (Chris Burge,
 http://www.cbs.dtu.dk/netgene/cbsnetgene.html), and NetPlantGene
 (http://www.cbs.dtu.dk/netgene/cbsnetgene.html), searches of the
 complete sequence against a peptide database and the Arabidopsis
 and Rice EST databases at TIGR, and the maize EST database at
 Genbank (http://www.tigr.org/tdb/at/at.html,
 http://www.tigr.org/tdb/home/tdb/cgi/index.html). Annotated genes
 are named to indicate the level of evidence for their annotation.
 Genes with similarity to other proteins are named after the
 database hits. Genes without significant peptide similarity but
 with EST similarity are named as 'unknown' proteins. Genes without
 protein or EST similarity, that are predicted by more than two gene
 prediction programs over most of their length are annotated as
 'hypothetical' proteins. Genes encoding tRNAs are predicted by
 tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
 Simple repeats are identified by repeatmasker (Arian Smit,
 http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of
 genomic sequence that are not annotated as genes but have predicted
 exons by GENE are annotated as misc features.
 FEATURES
 source Location/Qualifiers
 1..182756
 /organism="Oryza sativa"
 /cultivar="Nipponbare"
 /sub_species="japonica"
 /db_xref="taxon:4530"
 /clone="OSJNBA0049B20"
 complement(join(<179..204,660..775,896..1020,1264..1303,
 1587..2987,3531..3610,3722..3766))
 /gene="OSJNBA0049B20.1"
 complement(<179..3766)
 /gene="OSJNBA0049B20.1"
 /note="predicted by genscan and genefinder"

JOURNAL J. Gen. Microbiol. 135 (Pt 6), 1515-1520 (1989)
MEDLINE 90132571

FEATURES
Source Location/Qualifiers
1..1073
/organism="Bordetella pertussis"
/insertion_seq="18481 homolog"
/db_xref="taxon:520"

BASE COUNT 201 a 385 c 305 g 182 t

ORIGIN
Query Match 80.9%; Score 17.8; DB 1; Length 1073;
Best Local Similarity 90.5%; Pred. No. 5.7e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ggggacgagctcgtcgagg 22
||||| | ||||| |||||
Db 685 GGGGAAGCGCTCGCGGGG 665

Search completed: August 10, 2002, 02:58:05
Job time: 15691 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:21:46 ; Search time 1145.36 Seconds
(without alignments)
32.978 Million cell updates/sec

Title: US-09-672-126-11
Perfect score: 22
Sequence: 1 gggggacgagctgcggggg 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	22	100.0	22	AAF98741	Human IFN-alpha im
2	22	100.0	22	AAF99785	Immunostimulatory
3	18.8	85.5	22	AAF98739	Human IFN-alpha im
4	18.8	85.5	22	AAF98740	Human IFN-alpha im
5	18.8	85.5	22	AAF99783	Immunostimulatory
6	18.8	85.5	22	AAF99784	Immunostimulatory
7	17.8	80.9	16235	AAK86192	Human immune/haema
c 8	17.8	80.9	38186	AAZ32028	Human METH1 relate
c 9	17.8	80.9	38186	22 AAC90085	AC004449 cDNA clon

c 10	17.2	78.2	725	23	AA578214	DNA encoding novel
c 11	17.2	78.2	826	23	AA570398	DNA encoding novel
c 12	17.2	78.2	990	23	AA577182	DNA encoding novel
c 13	17.2	78.2	1023	23	AA584185	DNA encoding novel
c 14	17.2	78.2	1112	23	AA564267	DNA encoding novel
c 15	17.2	78.2	1283	23	AA564825	DNA encoding novel
c 16	17.2	78.2	2061	22	AA586409	4-amino-4-deoxycho
c 17	17.2	78.2	2616	23	AA590244	DNA encoding novel
c 18	17.2	78.2	3178	20	AA523728	WO9902653 Seq ID 1
c 19	17.2	78.2	3822	23	AA591811	DNA encoding novel
c 20	16.8	76.4	1239	14	AAQ61445	Lignin peroxidase
c 21	16.8	76.4	1273	13	AAQ31540	Lignin peroxidase
c 22	16.8	76.4	1350	23	AAI99917	Human alpha-2AAR e
c 23	16.8	76.4	1350	23	AAI99918	Human alpha-2AAR v
c 24	16.8	76.4	1666	14	AAQ61443	Lignin peroxidase
c 25	16.8	76.4	1810	14	AAQ51010	Lignin peroxidase
c 26	16.8	76.4	1810	14	AAQ61444	Lignin peroxidase
c 27	16.8	76.4	1918	13	AAQ31539	Lignin peroxidase
c 28	16.8	76.4	3269	22	AA575974	Human frizzled fam
c 29	16.8	76.4	7353	24	ABL32072	Human immune syste
c 30	16.2	73.6	21	22	AA598767	Human IFN-alpha im
c 31	16.2	73.6	21	22	AA598767	Immunostimulatory
c 32	16.2	73.6	276	20	AA587278	EST clone B0538.
c 33	16.2	73.6	308	21	AAQ3250	Human secreted pro
c 34	16.2	73.6	330	22	AAQ13792	Human breast cance
c 35	16.2	73.6	351	22	AA584578	Corn magnesium che
c 36	16.2	73.6	369	21	AAQ01688	Human secreted pro
c 37	16.2	73.6	400	22	ABA08414	Human secreted pro
c 38	16.2	73.6	428	19	AA530924	Human secreted pro
c 39	16.2	73.6	428	22	AA598403	Human CDNA clone B
c 40	16.2	73.6	716	22	AA592350	Human CDNA 5'-end
c 41	16.2	73.6	716	22	AA593869	Human CDNA clone r
c 42	16.2	73.6	726	22	AAH04520	Human CDNA clone (
c 43	16.2	73.6	784	22	AAH08823	Human CDNA clone (
c 44	16.2	73.6	831	21	AA507784	Fusarium venenatum
c 45	16.2	73.6	907	21	AA576631	Human ORFX ORF2186

ALIGNMENTS

RESULT 1
AAF98741
ID AAF98741 standard; DNA; 22 BP.
AC AAF98741;
AC AAF98741;
XX
XX 11-JUN-2001 (first entry)
DT
DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 11.
DE
DE Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
XX Synthetic.

OS
XX
XX Key Location/Qualifiers
FH modified_base 1..2
FT /*tag= a
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
FT modified_base 17..21
FT /*tag= b
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
XX
XX WO200122990-A2.
XX
XX 05-APR-2001.
XX
XX 27-SEP-2000; 2000WO-US26527.
XX
XX 27-SEP-1999; 99US-0156147.

XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX Hartmann G, Bratzler RL, Krieg A;
 XX WPI; 2001-290487/30.
 DR
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 XX
 PS Claim 201; Page 103; 168pp; English.
 XX
 CC The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX
 SQ Sequence 22 BP; 2 A; 4 C; 14 G; 2 T; 0 other;
 Query Match 100.0%; Score 22; DB 22; Length 22;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 gggggacgagctcgtcgggggg 22
 Db 1 gggggacgagctcgtcgggggg 22
 RESULT 2
 AAF99785
 ID AAF99785 standard; DNA; 22 BP.
 XX
 AC AAF99785;
 XX
 DT 12-JUN-2001 (first entry)
 DE
 DE Immunostimulatory nucleic acid #901.
 XX
 KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 OS Synthetic.
 XX
 PN WO200122972-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 25-SEP-2000; 2000WO-US26383.
 XX
 PR 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX
 PA (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GNBH.
 XX
 PI Krieg AM, Schetter C, Vollmer J;
 XX
 DR WPI; 2001-273485/28.
 XX
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 XX
 PS Claim 101; Page 57; 338pp; English.

XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 XX
 SQ Sequence 22 BP; 2 A; 4 C; 14 G; 2 T; 0 other;
 Query Match 100.0%; Score 22; DB 22; Length 22;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 gggggacgagctcgtcgggggg 22
 Db 1 gggggacgagctcgtcgggggg 22
 RESULT 3
 AAF98739
 ID AAF98739 standard; DNA; 22 BP.
 XX
 AC AAF98739;
 XX
 DT 11-JUN-2001 (first entry)
 DE
 DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 9.
 XX
 KW Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
 KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT modified_base 1..2
 FT /tag= a
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 FT modified_base 17..21
 FT /tag= b
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 XX
 PN WO200122990-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 27-SEP-2000; 2000WO-US26527.
 XX
 PR 27-SEP-1999; 99US-0156147.
 XX
 PA (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX
 PI Hartmann G, Bratzler RL, Krieg A;
 XX
 DR WPI; 2001-290487/30.
 XX
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 XX
 PS Claim 201; Page 103; 168pp; English.

XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX Sequence 22 BP; 3 A; 3 C; 13 G; 3 T; 0 other;

Query Match 85.5%; Score 18.8; DB 22; Length 22;
 Best Local Similarity 90.9%; Pred. No. 1e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgagctgcgtcggggg 22
 Db 1 gggggacgagctgcgtcggggg 22
 ||||| ||||| ||||| |||||

RESULT 4
 AAF98740
 ID AAF98740 standard; DNA; 22 BP.
 XX
 AC AAF98740;
 XX
 DT 11-JUN-2001 (first entry)
 XX
 DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 10.
 XX
 KW Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
 KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
 XX
 OS Synthetic.

Key Location/Qualifiers
 modified_base 1..2
 /tag= a
 /mod_base= "OTHER"
 /note= "phosphorothioate linkage"
 modified_base 17..21
 /tag= b
 /mod_base= "OTHER"
 /note= "phosphorothioate linkage"
 WO200122990-A2.

05-APR-2001.
 27-SEP-2000; 2000WO-US26527.
 27-SEP-1999; 99US-0156147.
 (COLE-) COLEY PHARM GROUP INC.
 (IOWA) UNIV IOWA RES FOUND.
 Hartmann G, Bratzler RL, Krieg A;
 WPI; 2001-290487/30.

Improving the efficacy of treatments involving the administration of
 interferon-alpha by co-administering an isolated immunostimulatory
 nucleic acid
 Claim 201; Page 103; 168pp; English.
 The present invention describes an improvement to a method requiring the
 administration of interferon alpha (IFN-alpha), involving administering
 an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 such nucleic acids are also provided. These may comprise oligonucleotides
 with phosphorothioate backbones, palindromes, or G-rich sequences. The

CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.

XX Sequence 22 BP; 2 A; 4 C; 14 G; 2 T; 0 other;
 SQ

Query Match 85.5%; Score 18.8; DB 22; Length 22;
 Best Local Similarity 90.9%; Pred. No. 1e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgagctgcgtcggggg 22
 Db 1 gggggacgagctgcgtcggggg 22
 ||||| ||||| ||||| |||||

RESULT 5
 AAF99783
 ID AAF99783 standard; DNA; 22 BP.
 XX
 AC AAF99783;
 XX
 DT 12-JUN-2001 (first entry)
 XX
 DE Immunostimulatory nucleic acid #899.
 XX
 KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX
 OS Synthetic.
 XX
 PN WO200122972-A2.
 XX
 PD 05-APR-2001.

25-SEP-2000; 2000WO-US26383.
 25-SEP-1999; 99US-0156113.
 27-SEP-1999; 99US-0156135.
 23-AUG-2000; 2000US-0227436.
 (IOWA) UNIV IOWA RES FOUND.
 (COLE-) COLEY PHARM GMBH.
 Krieg AM, Schetter C, Vollmer J;
 WPI; 2001-273485/28.
 Vaccinating against tumors, infectious diseases, allergies and asthma
 using immunostimulatory Py-rich and TG nucleic acids -
 Claim 101; Page 57; 338pp; English.

The present invention relates to a method for stimulating an immune
 response. The method comprises administering an immunostimulatory nucleic
 acid to a non-rodent subject in sufficient quantity to stimulate an
 immune response. The present sequence is one such immunostimulatory
 nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 haemophilus, campylobacter, clostridium, Escherichia coli and/or
 staphylococcus), fungal antigens and/or parasitic antigens. The method is
 also useful for preventing cancer, asthma, infectious disease, allergy or
 immune deficiency. The present sequence can also be used to redirect a
 Th2 to a Th1 immune response and to activate immune cells.
 Note: the present sequence may have a phosphorothioate backbone.

Sequence 22 BP; 3 A; 3 C; 13 G; 3 T; 0 other;

Query Match 85.5%; Score 18.8; DB 22; Length 22;
Best Local Similarity 90.9%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgacgtcgtcgggggg 22
| | | | | | | | | | | | | | | | | | | |
Db 1 gggggacgacgtcgtcgggggg 22

RESULT 6
AAF99784
ID AAF99784 standard; DNA; 22 BP.
XX AC AAF99784;
XX DT 12-JUN-2001 (first entry)
XX DE Immunostimulatory nucleic acid #900.

XX KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW Immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.

XX OS Synthetic.
XX PN WO200122972-A2.
XX PD 05-APR-2001.

XX PF 25-SEP-2000; 2000WO-US26383.
XX PR 25-SEP-1999; 99US-0156113.
XX PR 27-SEP-1999; 99US-0156135.
XX PR 23-AUG-2000; 2000US-0227436.

XX PA (TOWA) UNIV IOWA RES FOUND.
XX PA (COLE-) COLEY PHARM GMBH.

XX PI Krieg AM, Schetter C, Vollmer J;
XX DR WPI; 2001-273485/28.

XX PT Vaccinating against tumors, infectious diseases, allergies and asthma
XX PT using immunostimulatory py-rich and TG nucleic acids

XX PS Claim 101; Page 57; 338pp; English.

XX CC The present invention relates to a method for stimulating an immune
XX CC response. The method comprises administering an immunostimulatory nucleic
XX CC acid to a non-rodent subject in sufficient quantity to stimulate an
XX CC immune response. The present sequence is one such immunostimulatory
XX CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
XX CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
XX CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
XX CC and/or thymoxoviridae), bacterial antigens (e.g. toxoplasma,
XX CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
XX CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
XX CC also useful for preventing cancer, asthma, infectious disease, allergy or
XX CC immune deficiency. The present sequence can also be used to redirect a
XX CC Th2 to a Th1 immune response and to activate immune cells.
XX CC Note: the present sequence may have a phosphorothioate backbone.

XX SQ Sequence 22 BP; 2 A; 4 C; 14 G; 2 T; 0 other;

Query Match 85.5%; Score 18.8; DB 22; Length 22;
Best Local Similarity 90.9%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgacgtcgtcgggggg 22
| | | | | | | | | | | | | | | | | | | |
Db 1 gggggacgacgtcgtcgggggg 22

RESULT 7
AAK86192
ID AAK86192 standard; DNA; 16235 BP.
XX AC AAK86192;
XX DT 07-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:41004.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226688.
XX PR 22-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 06-SEP-2000; 2000US-0230437.
XX PR 06-SEP-2000; 2000US-0230438.
XX PR 08-SEP-2000; 2000US-0231242.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231244.
XX PR 08-SEP-2000; 2000US-0231413.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
DR useful for preventing, diagnosing and/or treating cancers and
XX metastasis.
XX Disclosure: SEQ ID NO 41004; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients' own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX Sequence 16235 BP; 2940 A; 4951 C; 4960 G; 3384 T; 0 other;
SQ
Query Match 80.9%; Score 17.8; DB 22; Length 16235;
Best Local Similarity 90.5%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 9999gacgagctcgctcg999g 21
Db 8934 9999gacgaggtggtcg999g 8954
RESULT 8
AAZ32028/c
ID AAZ32028 standard; DNA; 38186 BP.
XX
AC AAZ32028;
XX
DT 10-JAN-2000 (first entry)
XX Human METH1 related EST AC004449.
XX Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW angiogenesis inhibitor; abnormal wound healing; inflammation;
KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW diabetic retinopathy; macula degeneration; haemangioma; detection;
KW arterial-venous malformation; immune deficiency; ss.
XX

OS Homo sapiens.
 XX W09937660-A1.
 XX 29-JUL-1999.
 XX 22-JAN-1999; 99WO-US01313.
 XX 23-JAN-1998; 98US-0072298.
 XX 28-AUG-1998; 98US-0098539.
 XX (TRUE/) IRUELA-ARISPE L.
 PA (HAST/) HASTINGS G A.
 PA (RUBE/) RUBEN S M.
 XX IrueLa-Arispe L, Hastings GA, Ruben SM;
 XX WPI; 1999-590684/50.
 XX New isolated metalloprotease thrombospondin polypeptides, useful for
 XX treating hyperproliferative disorders, cancers or autoimmune disorders.
 XX
 XX Disclosure; Page 363-387; 457pp; English.
 XX AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
 CC metalloprotease thrombospondin (METH) proteins METH1 and METH2
 CC respectively. METH1 and METH2 have been found to be potent inhibitors of
 CC angiogenesis both in vitro and in vivo. They can be used for treating
 CC cancer and other disorders related to angiogenesis including abnormal
 CC wound healing, inflammation, rheumatoid arthritis, psoriasis,
 CC macula degeneration, haemangiomas, and arterial-venous malformations.
 CC They may be useful in treating deficiencies or disorders of the immune
 CC system, by activating or inhibiting the proliferation, differentiation,
 CC or mobilisation (chemotaxis) of immune cells. The etiology of these
 CC immune deficiencies or disorders may be genetic, somatic, such as
 CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
 CC toxins), or infectious. They can also be used to treat inflammatory
 CC conditions, both chronic and acute conditions. The products can also be
 CC used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to
 CC AAY49511 represent sequences given in the exemplification of the present
 CC invention.
 XX
 XX Sequence 38186 BP; 7571 A; 11503 C; 12193 G; 6919 T; 0 other;
 SQ
 Query Match 80.9%; Score 17.8; DB 20; Length 38186;
 Best Local Similarity 90.5%; Pred. No. 1.3e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 gggggacgagctcgtcggggg 21
 |||||
 Db 13585 GGGGACGAGGTGTCGGGGG 13565
 RESULT 9
 AAC90085/c
 ID AAC90085 standard; DNA; 38186 BP.
 XX AAC90085;
 AC AAC90085;
 XX 19-MAR-2001 (first entry)
 DT AC004449 cDNA clone.
 DE
 XX METH; metalloprotease; thrombospondin; angiogenesis inhibition;
 KW cancer therapy; benign tumour; ocular angiogenic disease;
 KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;
 KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
 KW scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
 KW coronary collateral; cerebral collateral; arteriovenous malformation;
 KW ischaemic limb angiogenesis; Osler-Webber syndrome; wound granulation;

KW plaque neovascularisation; telangiectasia; haemophilic joint; EST;
 KW angiofibroma; fibromuscular dysplasia; expressed sequence tag;
 KW Crohn's disease; atherosclerosis; birth control; ss.
 XX Unidentified.
 OS
 XX W0200071577-A1.
 PN
 XX 30-NOV-2000.
 PD
 XX 25-MAY-2000; 2000WO-US14462.
 PF
 XX 25-MAY-1999; 99US-0318208.
 PR 20-JUL-1999; 99US-0144882.
 PR 10-AUG-1999; 99US-0147823.
 PR 13-AUG-1999; 99US-0373658.
 PR 22-DEC-1999; 99US-0171503.
 PR 22-FEB-2000; 2000US-0183792.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
 PA (IRUELA-) IRUELA-ARISPE L.
 PA (HAST/) HASTINGS G A.
 PA (RUBE/) RUBEN S M.
 PA (JONA/) JONAK Z L.
 PA (TRUL/) TRULLI S H.
 PA (FORN/) FORNWALD J A.
 PA (TERRE/) TERRETT J A.
 XX IrueLa-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;
 PI Fornwald JA, Terrett JA;
 XX WPI; 2001-025136/03.
 XX METH1 and METH2 polynucleotides and encoded polypeptides, used to
 PT inhibit angiogenesis in the treatment of disorders such as cancer,
 PT rheumatoid arthritis and psoriasis -
 XX
 XX Claim 7; Pages 663-687; 768pp; English.
 PS The present invention relates to human METH1 and METH2, (ME for
 CC metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
 CC The present sequence is an expressed sequence tag (EST) for METH. METH
 CC can be used for inhibiting angiogenesis in an individual, and for
 CC treating cancer, benign tumours, an ocular angiogenic disease,
 CC rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
 CC vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
 CC scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
 CC coronary collaterals, cerebral collaterals, arteriovenous malformations,
 CC ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
 CC fibromuscular dysplasia, wound granulation, Crohn's disease or
 CC atherosclerosis. METH can also be used in birth control. METH can also
 CC be used in diagnostic methods for the prognosis of cancer.
 XX
 XX Sequence 38186 BP; 7571 A; 11503 C; 12194 G; 6919 T; 0 other;
 SQ
 Query Match 80.9%; Score 17.8; DB 22; Length 38186;
 Best Local Similarity 90.5%; Pred. No. 1.3e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 gggggacgagctcgtcggggg 21
 |||||
 Db 13585 GGGGACGAGGTGTCGGGGG 13565
 RESULT 10
 AAS78214/c
 ID AAS78214 standard; CDNA; 725 BP.
 XX
 AC AAS78214;

XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #14018.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI PI
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG14027.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 1; SEQ ID No 14018; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 725 BP; 176 A; 195 C; 207 G; 147 T; 0 other;

Query Match 78.2%; Score 17.2; DB 23; Length 725;
Best Local Similarity 86.4%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggggacgagctcgtcgggggg 22
||||| |||||||||
Db 531 GGGGGTGGAGCTCGTGGGGAG 510

RESULT 11
ID AAS70398/C
XX ID AAS70398 standard; cDNA; 826 BP.
XX AC AAS70398;
XX XX

DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #6202.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI PI
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG06211.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 1; SEQ ID No 6202; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 826 BP; 187 A; 227 C; 222 G; 190 T; 0 other;

Query Match 78.2%; Score 17.2; DB 23; Length 826;
Best Local Similarity 86.4%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggggacgagctcgtcgggggg 22
||||| |||||||||
Db 531 GGGGGTGGAGCTCGTGGGGAG 510

RESULT 12
ID AAS77182
XX ID AAS77182 standard; cDNA; 990 BP.
XX AC AAS77182;
XX XX
XX DT 13-FEB-2002 (first entry)

```
XX DE DNA encoding novel human diagnostic protein #12986.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG12995.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID No 12986; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 990 BP; 195 A; 299 C; 325 G; 171 T; 0 other;

Query Match 78.2%; Score 17.2; DB 23; Length 990;
Best Local Similarity 86.4%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggggacgagctcgtcggggg 22
Db 466 gggggcggagctcgtcggggag 487

RESULT 13
AAS84185
ID AAS84185 standard; cDNA; 1023 BP.
XX AC AAS84185;
XX XX
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #71.
```

```
DE DE DNA encoding novel human diagnostic protein #19989.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG19998.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID No 19989; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1023 BP; 212 A; 271 C; 271 G; 269 T; 0 other;

Query Match 78.2%; Score 17.2; DB 23; Length 1023;
Best Local Similarity 86.4%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggggacgagctcgtcggggg 22
Db 493 gggggcggagctcgtcggggag 514

RESULT 14
AAS64267
ID AAS64267 standard; cDNA; 1112 BP.
XX AC AAS64267;
XX XX
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #71.
```

```

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG00080.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID No 71; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1112 BP; 235 A; 325 C; 348 G; 204 T; 0 other;

Query Match 78.2%; Score 17.2; DB 23; Length 1112;
Best Local Similarity 86.4%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gggggacgagctcgtcgggggg 22
Db 583 gggggcggagctcgtcgggggg 604

RESULT 15
AAS64825/C
ID AAS64825 standard; cDNA; 1283 BP.
XX AC AAS64825;
XX AC
XX 13-FEB-2002 (first entry)
XX DE
XX DNA encoding novel human diagnostic protein #629.

```

```

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG000638.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID No 629; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1283 BP; 307 A; 350 C; 338 G; 288 T; 0 other;

Query Match 78.2%; Score 17.2; DB 23; Length 1283;
Best Local Similarity 86.4%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gggggacgagctcgtcgggggg 22
Db 531 gggggcggagctcgtcgggggg 510

Search completed: August 10, 2002, 03:21:48
Job time: 13679 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:11:08 ; Search time 9068.22 seconds
(without alignments)
32.744 Million cell updates/sec

Title: US-09-672-126-11
Perfect score: 22
Sequence: 1 ggggacgagctcgctggggg 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hic.*
- 9: gb_estl.*
- 10: gb_est2.*
- 11: gb_hic.*
- 12: gb_gss.*
- 13: em_gss_hum.*
- 14: em_gss_inv.*
- 15: em_gss_pla.*
- 16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.8	85.5	644	9	BB612044
2	17.8	80.9	280	10	BG597078
3	17.8	80.9	316	9	AU162819
4	17.8	80.9	326	10	BG127665
5	17.8	80.9	333	10	BG132128
6	17.8	80.9	358	10	D45979
7	17.8	80.9	361	9	AI483174
8	17.8	80.9	402	10	BG133866
9	17.8	80.9	406	10	BF098448
10	17.8	80.9	412	10	BF17067
11	17.8	80.9	424	9	AW442981
12	17.8	80.9	425	10	BI713967
13	17.8	80.9	439	9	AW037809
14	17.8	80.9	441	10	BF823810
15	17.8	80.9	485	10	BE450575
16	17.8	80.9	487	9	AW441770
17	17.8	80.9	489	9	AW441937

C 18	17.8	80.9	492	10	BG134116
C 19	17.8	80.9	495	9	AW455347
C 20	17.8	80.9	522	9	AW039330
C 21	17.8	80.9	554	9	AW649044
C 22	17.8	80.9	559	9	AW040811
C 23	17.8	80.9	560	10	BG135556
C 24	17.8	80.9	563	9	AW979358
C 25	17.8	80.9	582	9	AW399308
C 26	17.8	80.9	590	9	AW180748
C 27	17.8	80.9	590	9	AW622911
C 28	17.8	80.9	615	9	AW442342
C 29	17.8	80.9	630	10	BI960034
C 30	17.8	80.9	632	9	AW932234
C 31	17.8	80.9	641	9	AW398135
C 32	17.8	80.9	647	10	BG126836
C 33	17.8	80.9	653	10	BI954178
C 34	17.8	80.9	660	9	AW180075
C 35	17.8	80.9	720	10	BG125797
C 36	17.8	80.9	721	10	BG643338
C 37	17.8	80.9	730	10	BG123838
C 38	17.8	80.9	1077	12	AZ681805
C 39	17.8	80.9	1205	10	BF312709
C 40	17.4	79.1	559	10	BM491336
C 41	17.4	79.1	676	12	BH559375
C 42	17.2	78.2	322	9	BB252413
C 43	17.2	78.2	429	10	BE604929
C 44	17.2	78.2	434	10	BG907547
C 45	17.2	78.2	503	9	AI987353

ALIGNMENTS

RESULT 1

BB612044	BB612044	musculus	644 bp	mRNA	linear	EST 26-OCT-2001
LOCUS	BB612044	RIKEN full-length enriched, 15 days embryo head Mus				
DEFINITION	BB612044	musculus	CDNA clone 4022436K02 5', mRNA sequence.			
ACCESSION	BB612044	BB612044.1	GI:16453123			
VERSION	BB612044	EST.				
KEYWORDS	BB612044	house mouse.				
SOURCE	BB612044	Mus musculus				
ORGANISM	BB612044	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	BB612044	1 (bases 1 to 644)				
AUTHORS	BB612044	Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.				
TITLE	BB612044	RIKEN Mouse ESTs (Arakawa, T., et al. 2001)				
JOURNAL	BB612044	Unpublished (2001)				
COMMENT	BB612044	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp URI: http://genome.gsc.riken.go.jp/ Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

FEATURES

source

Location/Qualifiers

1. .644
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4022436K02"
 /clone_lib="RIKEN full-length enriched, 15 days embryo head"
 /sex="mixed"
 /tissue_type="head"
 /dev_stage="15 days embryo"
 /lab_host="DH10B"
 /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken. Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGAGTATTAATTAATTCGCCGCCGCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC 1"

BASE COUNT 115 a 198 c 203 g 128 t
 ORIGIN

Query Match 85.5%; Score 18.8; DB 9; Length 644;
 Best Local Similarity 90.9%; Pred. No. 2.2e+03;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggagagctctgcggggg 22
 |||||
 Db 91 GGGGACGAGCGGCGGCGG 112

RESULT

BG597078/c

LOCUS 280 bp mRNA linear EST 12-APR-2001
 DEFINITION EST495756 cSTS Solanum tuberosum cDNA clone cSTS16B18 5' sequence, mRNA sequence.

ACCESSION BG597078

VERSION 1 GI:13615218

KEYWORDS EST.

SOURCE potato.

ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 280)
 AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chieming,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.

TITLE
JOURNAL

COMMENT

Generations of ESTs from sprouting potato eyes

Unpublished (2000)

Contact: Cathy Ronning

The Institute for Genomic Research

For clone info: please contact Research Genetics, Libraries

Division tel 1-800-711-6195, email cdna@resgen.com

Seq primer: M13F-R.

FEATURES

source

Location/Qualifiers

1. .280
 /organism="Solanum tuberosum"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cSTS16B18"
 /clone_lib="cSTS"
 /tissue_type="sprouting eyes from tubers"
 /dev_stage="12-14 weeks post harvest"
 /lab_host="SOLR"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."

BASE COUNT 61 a 75 c 61 g 83 t
 ORIGIN

Query Match 80.9%; Score 17.8; DB 10; Length 280;

Best Local Similarity 90.5%; Pred. No. 4.6e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ggggacagctctgcggggg 22

|||||

Db 31 GGGGACGAGCTCTTCGGCGG 11

RESULT

AUI62819

LOCUS

DEFINITION AUI62819 Rice green shoot Oryza sativa cDNA clone S10902, mRNA sequence.

ACCESSION AUI62819

VERSION AUI62819.1 GI:11026218

KEYWORDS EST.

SOURCE Oryza sativa.

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 316)

AUTHORS Sasaki,T. and Yamamoto,K.

TITLE Rice cDNA from green shoot (2000)

JOURNAL

COMMENT

Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki

305-8602, Japan

Tel: 81-298-38-7441

Fax: 81-298-38-7468

Email: tsasaki@nri.affrc.go.jp, URL: <http://rgp.dna.affrc.go.jp/>

PROJECT "RGP",

S10902_12A.

FEATURES

source

Location/Qualifiers

1. .316
 /organism="Oryza sativa"
 /strain="Nipponbare"
 /db_xref="taxon:4530"
 /clone="S10902"
 /clone_lib="Rice green shoot"
 /note="Green shoot (8 days old)"

BASE COUNT 109 a 54 c 102 g 50 t 1 others
 ORIGIN

source

```

1..406
/organism="Lycopersicon esculentum"
/cultivar="TA492"
/db_xref="taxon:4081"
/clone="cLEW27C24"
/clone_lib="tomato nutrient deficient roots"
/tissue_type="roots"
/dev_stage="5-6 weeks old"
/lab_host="SOLR"
/note="Vector: pBluescriptSKCudapt; Site_1: 5' EcoRI; Site_2: 3' XhoI; Roots were harvested from plants grown under the following deficiencies/stresses: 10 mM Al, Zn, P, K, Fe, N. mRNA was isolated from individual treatments. Proportional aliquots of mRNA of each treatment were mixed and used for library construction."
BASE COUNT      91 a  97 c 100 g 118 t
ORIGIN
Query Match      80.9%; Score 17.8; DB 10; Length 406;
Best Local Similarity 90.5%; Pred. No. 4.8e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ggggacgagctcgctggggg 22
|| ||||| ||||| |||
Db 67 GGCAGCAGCTCGTCGGCGG 47

RESULT 10
BF1717067/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

USDA-WashU Neospora EST Project
Contact: Sandy Clifton, Ph.D. - Neospora
USDA-WashU Neospora EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Contact David Sibley (toxoe@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability.
Seq primer: -40RP from Gibco
High quality sequence stop: 233.
Location/Qualifiers
1..412
/organism="Neospora caninum"
/strain="Nc-1"
/db_xref="taxon:29176"
/clone_lib="Nc 1314 Tachyzoite cDNA"
/dev_stage="Tachyzoite"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; This library was constructed by Steve Fogarty, Robert Cole, and Keliang Tang at Washington University. cDNAs were synthesized from poly(A)+ RNA by oligo d(T) priming, size-selected and directionally cloned into the

```

```

Uni-ZAP XR lambda vector (Stratagene). The primary library was mass excised as phagemids and rescued in SOLR cells. The plasmid library was recovered from the SOLR cells and transformed in mass into DH10B (GeneHog, Research Genetics, Inc.) for sequencing. WARNING: the library may contain a small percentage of contaminants from human fibroblast cells."
BASE COUNT      108 a  96 c 123 g  84 t  1 others
ORIGIN
Query Match      80.9%; Score 17.8; DB 10; Length 412;
Best Local Similarity 90.5%; Pred. No. 4.8e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgagctcgctggggg 21
||| ||||| ||||| |||
Db 83 GGGTGACGAGCTCGTCGAGG 63

RESULT 11
AW442981/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

Generation of ESTs from tomato callus (mixed elicitor)
Unpublished (1999)
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1..424
/organism="Lycopersicon esculentum"
/cultivar="Rio Grande Ptor"
/db_xref="taxon:4081"
/clone="cLET42N19"
/clone_lib="tomato mixed elicitor, BTI"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="XL1-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLET - inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. ECORI site was destroyed during cloning."
BASE COUNT      93 a 100 c  98 g 133 t
ORIGIN
Query Match      80.9%; Score 17.8; DB 9; Length 424;
Best Local Similarity 90.5%; Pred. No. 4.9e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ggggacgagctcgctggggg 22
|| ||||| ||||| |||

```

Db 91 GCGACGAGCTCGCGGGG 71

RESULT 12

BI713967/c

LOCUS

DEFINITION

1c87h08.x1 Melton Normalized Mixed Mouse Pancreas 1 NI-MMS1 Mus

musculus cDNA 3' similar to TR:064410 O64410 CYTOCHROME P450

MONOOXYGENASE ; , mRNA sequence.

BI713967

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..425

/organism="Mus musculus"

/strain="ICR"

/db_xref="taxon:10090"

/clone_lib="Melton Normalized Mixed Mouse Pancreas 1

NI-MMS1"

/sex="Both for embryonic & newborn, male for adult and

adult islet"

/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,

adult, mixed"

/lab_host="DH10B"

/note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; Five

libraries representing E10.5/12.5 pancreatic bud, E16.5

pancreas, newborn pancreas, adult pancreas, and adult

islets of Langerhans were separately constructed using

Superscript Plasmid Library Kit (Life Technologies). cDNA

was made by oligo-dt priming and size-selected by column

fractionation. Libraries were amplified once on solid

support and plasmid DNA from each library was prepared

and mixed in equal amounts. The mixed library DNA was

normalized by method #4 from Bonaldo, Lennon, and Soares

1996 Genome Research 6:791-806; 0.5 microgram

single-stranded mixed library plasmid DNA was mixed with

5 micrograms PCR product representing mixed library

inserts and hybridized to an EcoT of 6. Single-stranded

(unhybridized) plasmids were isolated by hydroxyapatite

chromatography and used to make this library."

73 a 136 c 129 g 87 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

19; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

Score 17.8; DB 10; Length 425;

Pred. No. 4.9e+03;

0; Mismatches

2; Indels

0; Gaps

0;

80.9%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

QY 1 gggggacgagctcgctggggg 21

|||||

Db 161 GGGGAGAGCGCGCGGGG 141

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

QY 13

AW037809/c

LOCUS

DEFINITION

EST279438 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA

clone cLET3D8, mRNA sequence.

AW037809

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..439

/organism="Lycopersicon esculentum"

/cultivar="Rio Grande Ptor"

/db_xref="taxon:4081"

/clone="cLET3D8"

/clone_lib="tomato mixed elicitor, BTI"

/tissue_type="leaf"

/dev_stage="4-6 week old plants"

/lab_host="XLI-Blue MRP"

/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:

XhoI; cLET - Inoculated with a variety of disease response

elicitors. Plants exposed to 2,6 dichloroisonicotinic

acid, BTH, jasmonic acid, ethylene, fenthion, EIX,

okadaic acid, or systemin prior to tissue harvest. EcoRI

site was destroyed during cloning."

97 a 98 c 108 g 135 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

19; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

80.9%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

90.5%;

QY 2 ggggacgagctcgctggggg 22

|||||

Db 133 GCGACGAGCTCGCGGGG 113

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1225
; TELEFAX: 612-305-1228
;
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1598 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: DNA (genomic)
US-09-155-036-18

```

```
Query Match      71.0%; Score 14.2; DB 4; Length 1598;
Best Local Similarity 84.2%; pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

Qy 2 gggacgatcgttggggg 20
|||
Db 1200 GCGACGATCGATGGGCG 1218

RESULT 12
US-09-155-036-17
; Sequence 17, Application US/09155036
; Patent No. 6265201
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF MINNESOTA
; TITLE OF INVENTION: DNA MOLECULES AND PROTEIN DISPLAYING
; TITLE OF INVENTION: IMPROVED TRIAZINE COMPOUND DEGRADING ABILITY
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:

```
Query Match      71.0%; Score 14.2; DB 4; Length 1633;
Best Local Similarity 84.2%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

2y 2 ggggacgacgcttggggggg 20

Dbb 1235 GGCACGATCGATGGGCG 1253

```

RESULT 13
US-09-155-036-21
; Sequence 21, Application US/09155036
; Patent No. 6265201
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF MINNESOTA
; TITLE OF INVENTION: DNA MOLECULES AND PROTEIN DISPLAYING
; TITLE OF INVENTION: IMPROVED TRIAZINE COMPOUND DEGRADING ABILITY
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
; STREET: 119 No. 6265201th Fourth Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55401

```

Query Match	71.0%	Score 14.2;	DB 4;	Length 1674;
Best Local Similarity	84.2%;	Pred. NO. 1.4e+02;		
Matches 16;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

QY 2 ggggacgatcgtggggg 20
|||
db 1276 GCGACGATCGATGGGGG 1294

RESULT 14
US-09-362-473-3/c
; Sequence 3, Application US/09362473
; Patent No. 6218169
GENERAL INVENTION:
; APPLICANT: Cahoon, Edgar B.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Morgante, Michele
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
TITLE OF INVENTION: Aromatic Amino Acid Catabolism Enzymes
FILE REFERENCE: BB-1197
CURRENT APPLICATION NUMBER: US/09/362,473
CURRENT FILING DATE: 1999-07-28

TREATM

;; FILING DATE: 60/004914
;; APPLICATION NUMBER: 60/004914
;; FILING DATE: OCTOBER 6, 1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FLOYD, LINDA A.
;; REGISTRATION NUMBER: 33,692
;; REFERENCE/DOCKET NUMBER: CR-9677
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 302-892-8112
;; TELEFAX: 302-773-0164
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 384 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; STRAIN: P14K
;; US-09-103-434-21

Query Match 71.0%; Score 14.2; DB 3; Length 384;
Best Local Similarity 84.2%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ggggacgacgtgtgggggg 20
||||| ||||| ||||| |||||
Db 327 GGGGTCGATCGTGGCGGG 309

RESULT 7

US-09-687-594-21/c
; Sequence 21, Application US/09687594
; Patent No. 6251650

GENERAL INFORMATION:

;; APPLICANT: ROBERT D. FALLON
;; APPLICANT: MARK S. PAYNE
;; APPLICANT: MARK J. NELSON
;; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS ENCODING
;; TITLE OF INVENTION: STEREOSPECIFIC NITRILE HYDRATASE AND AMIDASE ENZYMES AND
;; TITLE OF INVENTION: RECOMBINANT ORGANISMS EXPRESSING THOSE ENZYMES USEFUL FOR
;; TITLE OF INVENTION: THE PRODUCTION OF CHIRAL AMIDES AND ACIDS
;; NUMBER OF SEQUENCES: 28
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
;; STREET: 1007 MARKET STREET
;; CITY: WILMINGTON
;; STATE: DELAWARE
;; COUNTRY: UNITED STATES OF AMERICA
;; ZIP: 19898

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: FLOPPY DISK
;; COMPUTER: IBM PC COMPATIBLE
;; OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
;; SOFTWARE: MICROSOFT WORD 2.0C
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/687,594
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/726,136
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FLOYD, LINDA A.
;; REGISTRATION NUMBER: 33,692
;; REFERENCE/DOCKET NUMBER: CR-9677
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 302-892-8112
;; TELEFAX: 302-773-0164
;; INFORMATION FOR SEQ ID NO: 21:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 384 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; STRAIN: P14K
;; US-09-687-594-21

Query Match 71.0%; Score 14.2; DB 4; Length 384;
Best Local Similarity 84.2%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ggggacgacgtgtgggggg 20
||||| ||||| ||||| |||||
Db 327 GGGGTCGATCGTGGCGGG 309

RESULT 8

US-08-818-112-138/c
; Sequence 138, Application US/08818112
; Patent No. 6290969

GENERAL INFORMATION:

;; APPLICANT: Reed, Steven G.
;; APPLICANT: Skeiky, Yasir A.W.
;; APPLICANT: Dillon, Davin C.
;; APPLICANT: Campos-Neto, Antonio
;; APPLICANT: Houghton, Raymond
;; APPLICANT: Vedvick, Thomas S.
;; APPLICANT: Twardzik, Daniel R.
;; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
;; NUMBER OF SEQUENCES: 153
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: SEED and BERRY LLP
;; STREET: 6300 Columbia Center, 701 Fifth Avenue
;; CITY: Seattle
;; STATE: Washington
;; COUNTRY: USA
;; ZIP: 98104-7092

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/818,112
;; FILING DATE: 13-MAR-1997
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Maki, David J.
;; REGISTRATION NUMBER: 31,392
;; REFERENCE/DOCKET NUMBER: 210121.411C6
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 622-4900
;; TELEFAX: (206) 682-6031
;; INFORMATION FOR SEQ ID NO: 138:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 882 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; US-08-818-112-138

Query Match 71.0%; Score 14.2; DB 4; Length 882;
Best Local Similarity 84.2%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

US-09-155-036-9

; Sequence 9, Application US/09155036
; Patent No. 6265201
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF MINNESOTA
; TITLE OF INVENTION: DNA MOLECULES AND PROTEIN DISPLAYING
; TITLE OF INVENTION: IMPROVED TRIAZINE COMPOUND DEGRADING ABILITY
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
; STREET: 119 No. 6265201th Fourth Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155.036
; FILING DATE: 16-JAN-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/035.404
; FILING DATE: 17-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCORMACK, MYRA M.
; REGISTRATION NUMBER: 36,602
; REFERENCE/DOCKET NUMBER: 110.00400201
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1225
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-155-036-9

Query Match 71.0%; Score 14.2; DB 4; Length 360;
Best Local Similarity 84.2%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ggggacgatcgttggggg 20
||| ||||| ||||| |||||
Db 208 GCGGACGATCGTGGGCG 226

RESULT 5

US-08-726-136-21/c
; Sequence 21, Application US/08726136
; Patent No. 5811286
; GENERAL INFORMATION:
; APPLICANT: ROBERT D. FALLON
; APPLICANT: MARK S. PAYNE
; APPLICANT: MARK J. NELSON
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS ENCODING
; TITLE OF INVENTION: STEREOSPECIFIC NITRILE HYDRATASE AND AMIDASE ENZYMES AND
; TITLE OF INVENTION: RECOMBINANT ORGANISMS EXPRESSING THOSE ENZYMES USEFUL FOR
; TITLE OF INVENTION: THE PRODUCTION OF CHIRAL AMIDES AND ACIDS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898

; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
; SOFTWARE: MICROSOFT WORD 2.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,136
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/004914
; FILING DATE: OCTOBER 6, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA A.
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9677
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; STRAIN: P14K
US-08-726-136-21

Query Match 71.0%; Score 14.2; DB 1; Length 384;
Best Local Similarity 84.2%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ggggacgatcgttggggg 20
||| ||||| ||||| |||||
Db 327 GGGTCGATCGCTGGCGG 309

RESULT 6

US-09-103-434-21/c
; Sequence 21, Application US/09103434
; Patent No. 6133421
; GENERAL INFORMATION:
; APPLICANT: ROBERT D. FALLON
; APPLICANT: MARK S. PAYNE
; APPLICANT: MARK J. NELSON
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS ENCODING
; TITLE OF INVENTION: STEREOSPECIFIC NITRILE HYDRATASE AND AMIDASE ENZYMES AND
; TITLE OF INVENTION: RECOMBINANT ORGANISMS EXPRESSING THOSE ENZYMES USEFUL FOR
; TITLE OF INVENTION: THE PRODUCTION OF CHIRAL AMIDES AND ACIDS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
; SOFTWARE: MICROSOFT WORD 2.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/103,434
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,136

—

High quality sequence stop: 163.

FEATURES

source
Location/Qualifiers
1. .1025
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5201825"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site.1: NotI; Site.2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 221 a 280 c 242 g 281 t 1 others
ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 1025;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gggggacgacgtcgttggggg 20
||||| ||||| ||||| |||||
DB 711 GGGGGACCATCGTGTGGGG 730

RESULT 15

AGI136593/C
LOCUS AGI136593 1030 bp DNA linear GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-150D13.F, genomic survey sequence.
ACCESSION AGI136593
VERSION AGI136593.1 GI:16666271
KEYWORDS GSS: GSS (genome survey sequence).
SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
BAC Library clone:PTB-150D13.F.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (sites)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 1030)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB this BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. .1030
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-150D13.F"
/sex="male"
/cell_type="lymphoblast"

FEATURES

source
Location/Qualifiers
1. .1030
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-150D13.F"
/sex="male"
/cell_type="lymphoblast"

BASE COUNT 239 a 354 c 207 g 198 t 32 others
ORIGIN

Query Match 84.0%; Score 16.8; DB 12; Length 1030;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gggggacgacgtcgttggggg 20
||||| ||||| ||||| |||||
DB 755 GGGGGAGGATCGTGGGGGG 736

Search completed: August 10, 2002, 02:11:14
Job time: 13135 sec

```
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgacgtgtggggg 20
|||||
Db 377 GGGGACGATCGTGGGAGG 358

RESULT 12
BF244081/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 902)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM941 row: h column: 02
High quality sequence stop: 557.
Location/Qualifiers
1. .902
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4080817"
/clone_lib="NIH_MGC_57"
/tissue_type="glioblastoma"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: brain; Vector: pDNR-LIB (Clontech); Site:1:
SfiI (ggcgctcgcc); Site:2: SfiI (ggcattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATCTAGAGCGGCGGCGGACATG-dt(30)BN-3'
sequence: 5'-ATCTAGAGCGGCGGCGGCGGACATG-dt(30)BN-3'
(Where B = A, C, G, or T). Average
insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
1 others
BASE COUNT 213 a 234 c 215 g 239 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 902;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgacgtgtggggg 20
|||||
Db 849 GGGGACGACGCTGGGGG 830

RESULT 13
BM469584/c
LOCUS
DEFINITION
ACCESSION
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 1002)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM11505 row: l column: 18
```

```
BM469584.1 GI:18518626
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 1002)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM12265 row: n column: 19
High quality sequence stop: 518.
Location/Qualifiers
1. .1002
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5551698"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lymph; Vector: PCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: This is a NIH_MGC Library."
BASE COUNT 212 a 319 c 287 g 184 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 1002;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgacgtgtggggg 20
|||||
Db 869 GGGGAGATCCTTGGGGG 850

RESULT 14
BF524320
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 1025)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM11505 row: l column: 18
```



```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. .643
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone_lib="unpublished oligo-capped cDNA library, C."
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
BASE COUNT 222 a 85 c 153 g 179 t 4 others
ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 643;
Best Local Similarity 90.0%; Pred. No. 1.5e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgacgttggtggggg 20
||||| ||| |||||
Db 358 GGGGACGATGTTGGGGGG 377

RESULT 8
AG131350/c
LOCUS
DEFINITION
Pan troglodytes DNA, clone: PTB-143G07.R, genomic survey sequence.
ACCESSION
AG131350
VERSION
AG131350.1 GI:16661028
KEYWORDS
GSS: GSS (genome survey sequence).
SOURCE
Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
BAC Library clone:PTB-143G07.R.
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (sites)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
REFERENCE
2 (bases 1 to 645)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
TITLE
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpansegsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY

Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
FEATURES
Location/Qualifiers
1. .645
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-143G07.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT 154 a 295 c 100 g 92 t 4 others
ORIGIN

Query Match 84.0%; Score 16.8; DB 12; Length 645;
Best Local Similarity 90.0%; Pred. No. 1.5e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgacgttggtggggg 20
||||| ||| |||||
Db 550 GGGGACGATGTTGGGGGG 531

RESULT 9
BB654689
LOCUS
DEFINITION
700 bp mRNA linear EST 26-OCT-2001
cDNA clone D030035K01 5', mRNA sequence.
ACCESSION
BB654689
VERSION
BB654689.1 GI:16488517
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 700)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
Hiramatsu, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T.,
Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura,
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa

```


strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected, histologically normal prostate epithelial cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

BASE COUNT 139 a 122 c 121 g 101 t
ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 483;

Best Local Similarity 94.7%; Pred. No. 7.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggggacgacgtctggggg 19
||||| ||||||| |||||

Db 48 GGGGGCCGATCGTGGGG 30

RESULT 2

BG789447/c

LOCUS 6HRm46 6HR Nitrogen-limited Schizopyllum library Schizopyllum commune cDNA 5' similar to mannose-1-phosphate guanylyltransferase, mRNA sequence. EST 16-MAY-2001

ACCESSION BG789447.1 GI:14124998

VERSION BG789447

KEYWORDS Schizopyllum commune.

SOURCE Schizopyllum commune.

ORGANISM Schizopyllum commune.

REFERENCE Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes; Agaricales; Schizopyllaceae; Schizopyllum.

AUTHORS 1 (bases 1 to 492)

TITLE Guettler, S., Lucchese, S.A., Honaas, L.A., Hittinger, C.T., Green, A., Lilly, W.W., and Gathman, A.C.

JOURNAL More expressed sequence tags from Schizopyllum commune

COMMENT nitrogen-replete and nitrogen-limited libraries

Unpublished (2001)

Contact: Gathman AC

Biology Department

Southeast MO State University

1 University Plaza, Cape Girardeau, MO 63701, USA

Tel: 5736512361

Fax: 5739866433

Email: agathman@biology.smo.edu

Seq primer: T3

POLYA-NO.

Location/Qualifiers

1. .492

/organism="Schizopyllum commune"

/strain="4-40"

/db_xref="taxon:5334"

/clone_lib="6HR Nitrogen-limited Schizopyllum library"

/tissue_type="mycelium"

/note="Vector: lambda zap; Site 1: EcoRI; Site 2: XhoI; 4-day-old mycelia of Schizopyllum commune were transferred from minimal (nitrogen-replete) medium to low-nitrogen medium. RNA was extracted six hours after transfer and cDNAs prepared."

BASE COUNT 71 a 189 c 128 g 102 t 2 others
ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 492;

Best Local Similarity 90.0%; Pred. No. 1.4e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgacgtctggggg 20

||||| ||||| ||||| |||||

Db 35 GGGGGACGCTCGTGGGGG 16

RESULT 3

BJ134969

LOCUS BJ134969

DEFINITION BJ134969 unpublished oligo-capped cDNA library, C. elegans LI stage

Caenorhabditis elegans cDNA clone yk1094e12 3', mRNA sequence.

ACCESSION BJ134969

VERSION BJ134969.1 GI:18295126

KEYWORDS EST.

SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans.

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

Rhabditidae; Peloderinae; Caenorhabditis.

AUTHORS 1 (bases 1 to 567)

Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.

and Sugano, S.

TITLE A complementary view of the C.elegans genome

JOURNAL Unpublished (2002)

COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. 567

/organism="Caenorhabditis elegans"

/strain="N2"

/db_xref="taxon:6239"

/clone_lib="yk1094e12"

/cldn_lib="unpublished oligo-capped cDNA library, C.

elegans LI stage"

/sex="hermaphrodite"

/tissue_type="whole animal"

/dev_stage="L1"

BASE COUNT 200 a 70 c 143 g 153 t 1 others

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 567;

Best Local Similarity 90.0%; Pred. No. 1.4e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgacgtctggggg 20

||||| ||||| ||||| |||||

Db 349 GGGGGACGATCGTGGGGG 368

RESULT 4

FR0013681/c

LOCUS FR0013681

DEFINITION F.rubripes GSS sequence, clone 130115B11, genomic survey sequence.

ACCESSION AL004927

VERSION AL004927.1 GI:2450497

KEYWORDS GSS; genome survey sequence.

SOURCE Takifugu rubripes

ORGANISM Takifugu rubripes

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;

Tetraodontidae; Takifugu.

REFERENCE 1 (bases 1 to 577)

AUTHORS Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrانيا, Y.,

Williams, G., and Brenner, S.

TITLE Direct Submission

JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource

Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hmp.mrc.ac.uk

COMMENT Vector: pBluescript II KS

V.type: phagemid

PRIMER: KS

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:11:11 ; Search time 9068.22 Seconds
(without alignments)
29.768 Million cell updates/sec

Title: US-09-672-126-13
Perfect score: 20
Sequence: 1 gggggacgacgttggggggg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	17.4	87.0	483	9	AA578972
C 2	16.8	84.0	492	10	BG789447
C 3	16.8	84.0	567	10	BG134969
C 4	16.8	84.0	577	12	FR0013681
C 5	16.8	84.0	580	12	CNS04BH3
C 6	16.8	84.0	630	10	BE961423
C 7	16.8	84.0	643	10	BE140169
C 8	16.8	84.0	645	12	AG131350
C 9	16.8	84.0	700	9	BB654689
C 10	16.8	84.0	793	12	B09090
C 11	16.8	84.0	862	12	CNS03JEQ
C 12	16.8	84.0	902	10	BF244081
C 13	16.8	84.0	1002	10	BM469584
C 14	16.8	84.0	1025	10	BF524320
C 15	16.8	84.0	1030	12	AG136593
C 16	16.8	84.0	1270	10	BM463696
C 17	16.8	84.0	1427	10	BG167937

C 18	16.8	84.0	1670	10	BE743142
C 19	16.4	82.0	907	12	CNS04RVO
C 20	16.4	82.0	908	10	BG300534
C 21	16.4	82.0	925	12	CNS03AGT
C 22	15.8	79.0	228	4	BB714070
C 23	15.8	79.0	327	10	BM151998
C 24	15.8	79.0	353	10	W06297
C 25	15.8	79.0	403	10	R27000
C 26	15.8	79.0	419	10	BM189385
C 27	15.8	79.0	457	9	AT005476
C 28	15.8	79.0	461	9	AW000785
C 29	15.8	79.0	492	10	BM132747
C 30	15.8	79.0	555	10	BM189895
C 31	15.8	79.0	556	10	BM189811
C 32	15.8	79.0	628	10	BF815550
C 33	15.8	79.0	635	10	BF856125
C 34	15.8	79.0	655	10	BF839093
C 35	15.8	79.0	710	12	AG130020
C 36	15.8	79.0	735	12	AG101070
C 37	15.8	79.0	768	10	BE957914
C 38	15.8	79.0	775	10	BF868211
C 39	15.8	79.0	812	10	BF522926
C 40	15.8	79.0	859	10	BM008797
C 41	15.8	79.0	913	10	BF215864
C 42	15.8	79.0	961	10	BE959048
C 43	15.8	79.0	987	10	BM458923
C 44	15.8	79.0	1029	10	BM454632
C 45	15.8	79.0	1032	10	BF731305

ALIGNMENTS

RESULT 1

AA578972/c

LOCUS

nf26f07.s1

DEFINITION

similar to gb:L06505 60S RIBOSOMAL PROTEIN L12 (HUMAN);, mrna

ACCESSION

AA578972

VERSION

AA578972.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 483)

AUTHORS

NCI-CCAP

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuqui, M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
DNA Sequencing by: Genome Systems Inc., Greg Lennon, Ph.D.
Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 394.
Location/Qualifiers
1. 483
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:914917"
/clone_lib="NCI-CCAP_Prl"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st

FEATURES

source

THIS PAGE BLANK (USPTO)

CC an example of an immunostimulatory oligonucleotide.

XX Sequence 20 BP; 0 A; 2 C; 13 G; 5 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 20;

Best Local Similarity 90.0%; Pred. No. 42;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gggggacgacgtgtggggggg 20

Db 1 gggggcgtcgtgtggggggg 20

RESULT 15

AAF98876

ID AAF98876 standard; DNA; 20 BP.

XX

AC AAF98876;

XX

DT 11-JUN-2001 (first entry)

XX

DE Immunostimulatory nucleic acid assay control oligo SEQ ID NO: 157.

XX

KW Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;

XX

KW viral infection; phosphorothioate backbone; palindromic; cancer; ds.

XX

OS Synthetic.

XX

PH Key Location/Qualifiers

FT modified_base 1..2

FT /*tag= a

FT /mod_base= "OTHER"

FT /note= "phosphorothioate linkage"

FT modified_base 15..19

FT /*tag= b

FT /mod_base= "OTHER"

FT /note= "phosphorothioate linkage"

XX WO200122990-A2.

XX

PN 05-APR-2001.

XX

PD 27-SEP-2000; 2000WO-US26527.

XX

PF 27-SEP-1999; 99US-0156147.

XX

PR (COLE-) COLEY PHARM GROUP INC.

XX

PA (IOWA) UNIV IOWA RES FOUND.

XX

PI Hartmann G, Bratzler RL, Krieg A;

XX

DR WPI; 2001-290487/30.

XX

PT Improving the efficacy of treatments involving the administration of

XX

PT interferon-alpha by co-administering an isolated immunostimulatory

XX

PT nucleic acid -

XX

PS Example 17; Page 165; 168pp; English.

XX

CC The present invention describes an improvement to a method requiring the

XX

CC administration of interferon alpha (IFN-alpha), involving administering

XX

CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of

XX

CC such nucleic acids are also provided. These may comprise oligonucleotides

XX

CC with phosphorothioate backbones, palindromes, or G-rich sequences. The

XX

CC sequences of the invention are useful in the treatment of proliferative

XX

CC diseases, such as cancers, and viral infections. The present sequence is

XX

CC an example of an immunostimulatory oligonucleotide.

XX

XX Sequence 20 BP; 2 A; 0 C; 13 G; 5 T; 0 other;

Query Match

84.0%; Score 16.8; DB 22; Length 20;

Best Local Similarity 90.0%; Pred. No. 42;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gggggacgacgtgtggggggg 20

Db 1 gggggcgtcgtgtggggggg 20

Search completed: August 10, 2002, 03:21:50

Job time: 13681 sec

CC response. The method comprises administering an immunostimulatory nucleic acid to a non-rodent subject in sufficient quantity to stimulate an immune response. The present sequence is one such immunostimulatory nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma, haemophilus, campylobacter, clostridium, Escherichia coli and/or staphylococcus), fungal antigens and/or parasitic antigens. The method is also useful for preventing cancer, asthma, infectious disease, allergy or immune deficiency. The present sequence can also be used to redirect a Th2 to a Th1 immune response and to activate immune cells.

CC Note: the present sequence may have a phosphorothioate backbone.

XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

SQ

Query Match 87.0%; Score 17.4; DB 22; Length 20;
Best Local Similarity 94.7%; Pred. No. 21;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ggggacgacgtgttggggg 20
|||||
Db 1 ggggacgacgtgttggggg 19

RESULT 13

AAF98744
ID AAF98744 standard; DNA; 20 BP.

XX AAF98744;

XX 11-JUN-2001 (first entry)

DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 14.

XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindromic; cancer; ds.

XX Synthetic.

Key Location/Qualifiers
FH modified_base 1..2
FT /tag= a
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
FT modified_base 16..19
FT /tag= b
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"

XX WO200122990-A2.

XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26527.

XX 27-SEP-1999; 99US-0156147.

XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.

XX Hartmann G, Bratzler RL, Krieg A;

XX WPI; 2001-290487/30.

XX Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -

XX Claim 201; Page 103; 168pp; English.

XX The present invention describes an improvement to a method requiring the

CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.

XX Sequence 20 BP; 3 A; 3 C; 12 G; 2 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgacgtgttggggg 20
|||||
Db 1 ggggacgacgtgttggggg 20

RESULT 14

AAF98761
ID AAF98761 standard; DNA; 20 BP.

XX AAF98761;

XX 11-JUN-2001 (first entry)

XX Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 31.

XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindromic; cancer; ds.

XX Synthetic.

Key Location/Qualifiers
FH modified_base 1..2
FT /tag= a
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
FT modified_base 15..19
FT /tag= b
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"

XX WO200122990-A2.

XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26527.

XX 27-SEP-1999; 99US-0156147.

XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.

XX Hartmann G, Bratzler RL, Krieg A;

XX WPI; 2001-290487/30.

XX Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -

XX Claim 201; Page 103; 168pp; English.

XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is

PA (IOWA) UNIV IOWA RES FOUND.
 XX (COLE-) COLEY PHARM GMBH.
 PI Krieg AM, Schetter C, Vollmer J;
 XX WPI; 2001-273485/28.
 DR Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 XX Claim 101; Page 57; 338pp; English.
 PS
 CC The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy, or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 XX
 SQ Sequence 19 BP; 2 A; 3 C; 12 G; 2 T; 0 other;
 Query Match 87.0%; Score 17.4; DB 22; Length 19;
 Best Local Similarity 94.7%; Pred. No. 21;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ggggggacgcgttggggg 19
 DB 1 ggggggacgcgtcggggg 19
 RESULT 11
 AAF98765
 ID AAF98765 standard; DNA; 20 BP.
 XX
 AC AAF98765;
 XX
 DT 11-JUN-2001 (first entry)
 XX
 DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 35.
 XX
 KW Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
 KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT modified_base 1..2
 FT /*tag= a
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 FT modified_base 15..19
 FT /*tag= b
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 XX
 PN WO200122990-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 27-SEP-2000; 2000NO-US26527.
 XX
 PP 27-SEP-1999; 99US-0156147.
 XX
 PR 27-SEP-1999; 99US-0156147.
 XX
 PA (COLE-) COLEY PHARM GROUP INC.

PA (IOWA) UNIV IOWA RES FOUND.
 XX Hartmann G, Bratzler RL, Krieg A;
 PI WPI; 2001-290487/30.
 DR
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 XX Claim 201; Page 103; 168pp; English.
 PS
 CC The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX
 SQ Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;
 Query Match 87.0%; Score 17.4; DB 22; Length 20;
 Best Local Similarity 94.7%; Pred. No. 21;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 ggggacgcgttggggg 20
 DB 1 ggggacgcgtcggggg 19
 RESULT 12
 AAF99869
 ID AAF99869 standard; DNA; 20 BP.
 XX
 AC AAF99869;
 XX
 DT 12-JUN-2001 (first entry)
 XX
 DE Immunostimulatory nucleic acid #985.
 XX
 KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX
 OS Synthetic.
 XX
 PN WO200122972-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 25-SEP-2000; 2000NO-US26383.
 XX
 PR 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX
 PA (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 XX
 PI Krieg AM, Schetter C, Vollmer J;
 XX WPI; 2001-273485/28.
 DR
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 XX Claim 101; Page 59; 338pp; English.
 PS
 XX The present invention relates to a method for stimulating an immune

KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX Synthetic.
 OS
 XX WO200122972-A2.
 PN
 XX
 PD 05-APR-2001.
 XX
 XX 25-SEP-2000; 2000WO-US26383.
 PF
 XX 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX
 XX (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 XX
 XX Krieg AM, Schetter C, Vollmer J;
 PI
 XX WPI; 2001-273485/28.
 DR
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 XX using immunostimulatory Py-rich and TG nucleic acids -
 PT
 XX Claim 101; Page 58; 338pp; English.
 PS
 XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 XX
 XX Sequence 21 BP; 2 A; 3 C; 14 G; 2 T; 0 other;
 SQ
 Query Match 92.0%; Score 18.4; DB 22; Length 21;
 Best Local Similarity 95.0%; Pred. No. 7;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 gggggacgacgtctggggggg 20
 Db 1 gggggacgacgtctggggggg 20
 RESULT 9
 AAF99850
 ID AAF99850 standard; DNA; 20 BP.
 XX
 XX AAF99850;
 AC
 XX 12-JUN-2001 (first entry)
 DT
 XX Immunostimulatory nucleic acid #966.
 DE
 XX Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX
 OS Synthetic.
 XX
 XX WO200122972-A2.
 PN
 XX 05-APR-2001.
 PD

XX 25-SEP-2000; 2000WO-US26383.
 PF
 XX 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX
 XX (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 XX
 XX Krieg AM, Schetter C, Vollmer J;
 PI
 XX WPI; 2001-273485/28.
 DR
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 XX using immunostimulatory Py-rich and TG nucleic acids -
 PT
 XX Claim 101; Page 59; 338pp; English.
 PS
 XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 XX
 XX Sequence 20 BP; 2 A; 0 C; 13 G; 3 T; 2 other;
 SQ
 Query Match 90.0%; Score 18; DB 22; Length 20;
 Best Local Similarity 90.0%; Pred. No. 11;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 gggggacgacgtctggggggg 20
 Db 1 gggggacgacgtctggggggg 20
 RESULT 10
 AAF99754
 ID AAF99754 standard; DNA; 19 BP.
 XX
 XX AAF99754;
 AC
 XX 12-JUN-2001 (first entry)
 DT
 XX Immunostimulatory nucleic acid #870.
 DE
 XX Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX
 OS Synthetic.
 XX
 XX WO200122972-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 25-SEP-2000; 2000WO-US26383.
 PF
 XX 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX

```

RESULT 6
AAF98746
ID AAF98746 standard; DNA; 21 BP.
XX
AC AAF98746;
XX
DT 11-JUN-2001 (first entry)
XX
DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 16.
XX
KW Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..2
FT FT /*tag= a
FT FT /mod_base= "OTHER"
FT FT /note= "phosphorothioate linkage"
FT modified_base 16..20
FT FT /*tag= b
FT FT /mod_base= "OTHER"
FT FT /note= "phosphorothioate linkage"
XX
PN WO200122990-A2.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26527.
XX
PR 27-SEP-1999; 99US-0156147.
XX
PA (COLE-) COLEY PHARM GROUP INC.
PA (IOWA ) UNIV IOWA RES FOUND.
XX
PI Hartmann G, Bratzler RL, Krieg A;
XX
DR WPI; 2001-290487/30.
XX
PT Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
PS Claim 201; Page 103; 168pp; English.
XX
CC The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX
SQ Sequence 21 BP; 2 A; 3 C; 14 G; 2 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 21;
Best Local Similarity 95.0%; Pred. No. 7;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggggacgacgtgtggggg 20
   |||||
Db 1 gggggacgacgtgtggggg 20

RESULT 7
AAF99791
ID AAF99791 standard; DNA; 21 BP.
XX
AC AAF99791;
XX

```

```

DT 12-JUN-2001 (first entry)
XX
DE Immunostimulatory nucleic acid #907.
XX
KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS Synthetic.
XX
PN WO200122972-A2.
XX
PD 05-APR-2001.
XX
PF 25-SEP-2000; 2000WO-US26383.
XX
PR 25-SEP-1999; 99US-0156113.
PR 27-SEP-1999; 99US-0156135.
PR 23-AUG-2000; 2000US-0227436.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX
DR WPI; 2001-273485/28.
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 58; 338pp; English.
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SQ Sequence 21 BP; 2 A; 3 C; 14 G; 2 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 21;
Best Local Similarity 95.0%; Pred. No. 7;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggggacgacgtgtggggg 20
   |||||
Db 2 gggggacgacgtgtggggg 21

RESULT 8
AAF99792
ID AAF99792 standard; DNA; 21 BP.
XX
AC AAF99792;
XX
DT 12-JUN-2001 (first entry)
XX
DE Immunostimulatory nucleic acid #908.
XX
KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;

```

XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 6.9;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gggggacgacgtgtggggg 20
|||||
Db 1 gggggacgacgtgtggggg 20

RESULT 4

AAF98852
ID AAF98852 standard; DNA; 20 BP.

XX AAF98852;

DT 11-JUN-2001 (first entry)

XX Poly-G immunostimulatory nucleic acid SEQ ID NO: 133.

XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.

XX Synthetic.

PN WO200122990-A2.

XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26527.

XX 27-SEP-1999; 99US-0156147.

XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.

XX Hartmann G, Bratzler RL, Krieg A;

XX WPI; 2001-290487/30.

XX Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid.

PS Disclosure; Page 24; 168pp; English.

XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.

XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 6.9;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gggggacgacgtgtggggg 20
|||||
Db 1 gggggacgacgtgtggggg 20

RESULT 5

AAF98745

ID AAF98745 standard; DNA; 21 BP.

XX AAF98745;

DT 11-JUN-2001 (first entry)

XX Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 15.

XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.

XX Synthetic.

XX Key Location/Qualifiers
FT modified_base 1..2

FT /*tag= a
FT /mod_base= "OTHER"

FT /note= "phosphorothioate linkage"
FT modified_base 16..20

FT /*tag= b

FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"

PN WO200122990-A2.

XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26527.

XX 27-SEP-1999; 99US-0156147.

XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.

XX Hartmann G, Bratzler RL, Krieg A;

XX WPI; 2001-290487/30.

XX Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid.

PS Claim 201; Page 103; 168pp; English.

XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.

XX Sequence 21 BP; 2 A; 3 C; 14 G; 2 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 21;

Best Local Similarity 95.0%; Pred. No. 7;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gggggacgacgtgtggggg 20
|||||

Db 2 gggggacgacgtgtggggg 21
|||||

XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.
XX Hartmann G, Bratzler RL, Krieg A;
XX WPI; 2001-290487/30.
XX Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX Claim 201; Page 103; 168pp; English.
XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX Sequence 20 BP; 2 A; 2 C; 13 G; 3 T; 0 other;
SQ

Query Match 100.0%; Score 20; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggggacgacgtgtggggg 20
| | | | | | | | | | | | | | | | | |
Db 1 gggggacgacgtgtggggg 20

RESULT 2
AAF99789
ID AAF99789 standard; DNA; 20 BP.
XX AAF99789;
XX 12-JUN-2001 (first entry)
DE Immunostimulatory nucleic acid #905.
XX Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX Synthetic.
OS
XX WO200122972-A2.
XX 05-APR-2001.
XX 25-SEP-2000; 2000WO-US26383.
XX 25-SEP-1999; 99US-0156113.
PR 27-SEP-1999; 99US-0156135.
PR 23-AUG-2000; 2000US-0227436.
XX (IOWA) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX Krieg AM, Schetter C, Vollmer J;
XX WPI; 2001-273485/28.
DR
XX Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory py-rich and TG nucleic acids -
XX Claim 101; Page 58; 338pp; English.

XX The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX Sequence 20 BP; 2 A; 2 C; 13 G; 3 T; 0 other;
SQ

Query Match 100.0%; Score 20; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggggacgacgtgtggggg 20
| | | | | | | | | | | | | | | | | |
Db 1 gggggacgacgtgtggggg 20

RESULT 3
AAF98737
ID AAF98737 standard; DNA; 20 BP.
XX AAF98737;
XX 11-JUN-2001 (first entry)
DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 7.
XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX Synthetic.
OS
XX Key Location/Qualifiers
FH modified_base 1..2 /*tag= a
FT /*mod_base= "OTHER"
FT /*note= "phosphorothioate linkage"
FT modified_base 15..19 /*tag= b
FT /*mod_base= "OTHER"
FT /*note= "phosphorothioate linkage"
XX WO200122990-A2.
XX 05-APR-2001.
XX 27-SEP-2000; 2000WO-US26527.
XX 27-SEP-1999; 99US-0156147.
XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.
XX Hartmann G, Bratzler RL, Krieg A;
XX WPI; 2001-290487/30.
DR
XX Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX Claim 201; Page 103; 168pp; English.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:21:48 ; Search time 1145.36 Seconds
(without alignments)
29.980 Million cell updates/sec

Title: US-09-672-126-13

Perfect score: 20

Sequence: 1 gggggacgacgttggggg 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

1:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	AAF98743	Human IFN-alpha im
2	20	100.0	20	AAF99789	Immunostimulatory
3	18.4	92.0	20	AAF98737	Human IFN-alpha im
4	18.4	92.0	20	AAF98852	Poly-G Immunostimu
5	18.4	92.0	21	AAF98745	Human IFN-alpha im
6	18.4	92.0	21	AAF98746	Human IFN-alpha im
7	18.4	92.0	21	AAF98791	Immunostimulatory
8	18.4	92.0	21	AAF99792	Immunostimulatory
9	18	90.0	20	AAF99850	Immunostimulatory

10	17.4	87.0	19	22	AAF99754	Immunostimulatory
11	17.4	87.0	20	22	AAF98765	Human IFN-alpha im
12	17.4	87.0	20	22	AAF99869	Immunostimulatory
13	16.8	84.0	20	22	AAF98744	Human IFN-alpha im
14	16.8	84.0	20	22	AAF98761	Human IFN-alpha im
15	16.8	84.0	20	22	AAF98876	Immunostimulatory
16	16.8	84.0	20	22	AAF99790	Immunostimulatory
17	16.8	84.0	20	22	AAF99848	Immunostimulatory
18	16.8	84.0	20	22	AAF99849	Immunostimulatory
19	16.4	82.0	19	22	AAF98762	Human IFN-alpha im
20	16.4	82.0	19	22	AAF99865	Immunostimulatory
21	16.4	82.0	20	22	AAF98877	Immunostimulatory
22	16.4	82.0	4705	23	ABL10184	Drosophila melanog
23	16.4	82.0	6895	23	ABL10188	Drosophila melanog
24	15.8	79.0	6089	24	ABL32702	Human immune syste
25	15.8	79.0	294528	24	ABA03041	Listeria monocytog
26	15.4	77.0	289	21	AAC23974	Human secreted pro
27	15.2	76.0	20	22	AAF98742	Human IFN-alpha im
28	15.2	76.0	20	22	AAF99786	Immunostimulatory
29	15.2	76.0	30	22	AAF98884	IFN-1 inducing cod
30	15.2	76.0	72	21	AAA03795	Streptavidin displ
31	15.2	76.0	699	21	AAF13496	Aspergillus oryzae
32	15.2	76.0	1104	22	AAAL25097	Human breast cance
33	15.2	76.0	1329	22	AAF81359	Quorum sensing con
34	15.2	76.0	2958	23	ABL26044	Drosophila melanog
35	15.2	76.0	6568	24	ABL32447	Human immune syste
36	15.2	76.0	7862	23	ABL07766	Drosophila melanog
37	15.2	76.0	11920	23	ABL21028	Drosophila melanog
38	15.2	76.0	80251	23	ABL16442	Drosophila melanog
39	15.2	76.0	80251	23	ABL16448	Drosophila melanog
40	15	75.0	826	22	AAI95547	Human neuroblastom
41	14.8	74.0	618	22	AAF65780	Novel human polynu
42	14.8	74.0	1050	22	AAAS31313	Human cDNA encodin
43	14.8	74.0	1083	21	AAAS33352	Arabidopsis thalia
44	14.8	74.0	2011	23	AAAS89364	DNA encoding novel
45	14.8	74.0	2265	23	AAAS92867	DNA encoding novel

ALIGNMENTS

RESULT 1

ID AAF98743 standard; DNA; 20 BP.

XX AAF98743;

AC AAF98743;

DT 11-JUN-2001 (first entry)

XX Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 13.

DE Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;

KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.

XX Synthetic.

OS Synthetic.

XX Key

FH modified_base

FT Location/Qualifiers

FT 1..2

FT /*tag= a

FT /mod_base= "OTHER"

FT /note= "phosphorothioate linkage"

FT modified_base

FT 16..19

FT /*tag= b

FT /mod_base= "OTHER"

FT /note= "phosphorothioate linkage"

FT WO200122990-A2.

XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26527.

XX 27-SEP-1999; 99US-0156147.

artificial sequence.
1 (bases 1 to 20)
Krieg, A.M., Schetter, C. and Vollmer, J.C.
Immunostimulatory nucleic acids
Patent: WO 0122972-A 996 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US); Coley Pharmaceutical
GmbH (DE)

FEATURES
Source
Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630" 2 t
BASE COUNT 3 a 3 c 12 g
ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 20;
Best Local Similarity 90.0%; Pred. No. 5.4e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 gggggacgacgttggtggggg 20
||||| ||||||| |||||||
Db 1 GGGGACGATCGTCGGGGG 20

RESULT 14
AX104862
LOCUS AX104862 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1054 from Patent WO0122972.
ACCESSION AX104862
VERSION AX104862.1 GI:13921059
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg, A.M., Schetter, C. and Vollmer, J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 1054 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US); Coley Pharmaceutical
GmbH (DE)

FEATURES
Source
Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630" 5 t
BASE COUNT 0 a 2 c 13 g
ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 20;
Best Local Similarity 90.0%; Pred. No. 5.4e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 gggggacgacgttggtggggg 20
||||| ||||||| |||||||
Db 1 GGGGTCGTCGTCGGGGG 20

RESULT 15
AX104863
LOCUS AX104863 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1055 from Patent WO0122972.
ACCESSION AX104863
VERSION AX104863.1 GI:13921060
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg, A.M., Schetter, C. and Vollmer, J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 1055 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US); Coley Pharmaceutical
GmbH (DE)

FEATURES
Source
Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630" 5 t
BASE COUNT 2 a 0 c 13 g
ORIGIN
Query Match 84.0%; Score 16.8; DB 6; Length 20;
Best Local Similarity 90.0%; Pred. No. 5.4e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 gggggacgacgttggtggggg 20
||||| ||||||| |||||||
Db 1 GGGGATGATGTTGGGGG 20
Search completed: August 10, 2002, 02:58:07
Job time: 15693 sec

```
artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 1056 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES             Location/Qualifiers
     source          1..20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     modified_base    7
     modified_base    11
BASE COUNT  2 a      0 c      13 g      3 t      2 others
ORIGIN
1 gggggacgacgttggtggggg 20
||||| ||| |||||
Db 1 GGGGGANGATNGTGGGGG 20

Query Match 90.0%; Score 18; DB 6; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 10
AX104767
LOCUS AX104767 19 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 959 from Patent WO0122972.
ACCESSION AX104767
VERSION AX104767.1 GI:13920964
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 19)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 959 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES             Location/Qualifiers
     source          1..19
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT  2 a      3 c      12 g      2 t
ORIGIN
1 gggggacgacgttggtgggg 19
||||| ||| |||||
Db 1 GGGGGACGATCGTCGGGG 19

Query Match 87.0%; Score 17.4; DB 6; Length 19;
Best Local Similarity 94.7%; Pred. No. 3e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 11
AX104883
LOCUS AX104883 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1075 from Patent WO0122972.
ACCESSION AX104883
VERSION AX104883.1 GI:13921080
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.

artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 1075 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES             Location/Qualifiers
     source          1..20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT  2 a      3 c      13 g      2 t
ORIGIN
1 gggggacgacgttggtgggg 20
||||| ||| |||||
Db 1 GGGGACGATCGTCGGGGG 19

Query Match 87.0%; Score 17.4; DB 6; Length 20;
Best Local Similarity 94.7%; Pred. No. 2.9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 12
AX105137
LOCUS AX105137 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 35 from Patent WO0122990.
ACCESSION AX105137
VERSION AX105137.1 GI:13921287
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 35 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1..20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1..2
                     /note="Backbone has phosphorothioate linkages."
     misc_feature    3..14
                     /note="Backbone has phosphodiester linkages."
     misc_feature    15..19
                     /note="Backbone has phosphorothioate linkages."
     misc_feature    20
                     /note="Backbone has phosphodiester linkages."
BASE COUNT  2 a      3 c      13 g      2 t
ORIGIN
1 gggggacgacgttggtgggg 20
||||| ||| |||||
Db 1 GGGGACGATCGTCGGGGG 19

Query Match 87.0%; Score 17.4; DB 6; Length 20;
Best Local Similarity 94.7%; Pred. No. 2.9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 13
AX104804
LOCUS AX104804 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 996 from Patent WO0122972.
ACCESSION AX104804
VERSION AX104804.1 GI:13921001
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
```

```
GmbH (DE)
FEATURES             Location/Qualifiers
     source           1..21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT          2 a      3 c      14 g      2 t
ORIGIN
Query Match          92.0%; Score 18.4; DB 6; Length 21;
Best Local Similarity 95.0%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 gggggacgacgtgtggggg 20
    |||||
Db 2 GGGGGACGATCGTCGGGGG 21

RESULT 6
LOCUS AX104806 21 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 998 from Patent WO0122972.
ACCESSION AX104806
VERSION AX104806.1 GI:13921003
KEYWORDS
SOURCE
ORGANISM synthetic construct.
          artificial sequence.
REFERENCE 1 (bases 1 to 21)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 998 05-APR-2001.
        UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
        GmbH (DE)
FEATURES             Location/Qualifiers
     source           1..21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT          2 a      3 c      14 g      2 t
ORIGIN
Query Match          92.0%; Score 18.4; DB 6; Length 21;
Best Local Similarity 95.0%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 gggggacgacgtgtggggg 20
    |||||
Db 1 GGGGGACGATCGTCGGGGG 20

RESULT 7
LOCUS AX105117 21 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 15 from Patent WO0122990.
ACCESSION AX105117
VERSION AX105117.1 GI:13921267
KEYWORDS
SOURCE
ORGANISM synthetic construct.
          synthetic construct.
          artificial sequence.
REFERENCE 1 (bases 1 to 21)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
        interferon
JOURNAL Patent: WO 0122990-A 15 05-APR-2001;
        Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
        FOUNDATION (US)
FEATURES             Location/Qualifiers
     source           1..21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT          2 a      3 c      14 g      2 t
ORIGIN
Query Match          92.0%; Score 18.4; DB 6; Length 21;
Best Local Similarity 95.0%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 gggggacgacgtgtggggg 20
    |||||
Db 1 GGGGGACGATCGTCGGGGG 20

RESULT 8
LOCUS AX105118 21 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 16 from Patent WO0122990.
ACCESSION AX105118
VERSION AX105118.1 GI:13921268
KEYWORDS
SOURCE
ORGANISM synthetic construct.
          synthetic construct.
          artificial sequence.
REFERENCE 1 (bases 1 to 21)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
        interferon
JOURNAL Patent: WO 0122990-A 16 05-APR-2001;
        Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
        FOUNDATION (US)
FEATURES             Location/Qualifiers
     source           1..21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic Oligonucleotide"
BASE COUNT          1..2
                     /note="Backbone has phosphorothioate linkages."
misc_feature         3..15
                     /note="Backbone has phosphorothioate linkages."
misc_feature         16..20
                     /note="Backbone has phosphorothioate linkages."
misc_feature         21
                     /note="Backbone has phosphorothioate linkages."
BASE COUNT          2 a      3 c      14 g      2 t
ORIGIN
Query Match          92.0%; Score 18.4; DB 6; Length 21;
Best Local Similarity 95.0%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 gggggacgacgtgtggggg 20
    |||||
Db 2 GGGGGACGATCGTCGGGGG 21

RESULT 9
LOCUS AX104864 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1056 from Patent WO0122972.
ACCESSION AX104864
VERSION AX104864.1 GI:13921061
KEYWORDS
SOURCE
ORGANISM synthetic construct.
          synthetic construct.
```

```
Query Match      100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggggacgacgttggtggggg 20
   |||||
Db 1 GGGGGACGATCGTGGGGGG 20

RESULT 2
AX105115
LOCUS AX105115 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 13 from Patent WO0122990.
ACCESSION AX105115
VERSION AX105115.1 GI:13921265
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 13 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
misc_feature 1..2
/note="Backbone has phosphorothioate linkages."
misc_feature 3..15
/note="Backbone has phosphodiester linkages."
misc_feature 16..19
/note="Backbone has phosphorothioate linkages."
misc_feature 20
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 2 c 13 g 3 t
ORIGIN
1
//

Query Match      100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggggacgacgttggtggggg 20
   |||||
Db 1 GGGGGACGATCGTGGGGGG 20

RESULT 2
AX105115
LOCUS AX105115 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 13 from Patent WO0122990.
ACCESSION AX105115
VERSION AX105115.1 GI:13921265
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 13 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
misc_feature 1..2
/note="Backbone has phosphorothioate linkages."
misc_feature 3..15
/note="Backbone has phosphodiester linkages."
misc_feature 16..19
/note="Backbone has phosphorothioate linkages."
misc_feature 20
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 2 c 13 g 3 t
ORIGIN
1
//

Query Match      100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggggacgacgttggtggggg 20
   |||||
Db 1 GGGGGACGATCGTGGGGGG 20

RESULT 3
AX105109
LOCUS AX105109 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 7 from Patent WO0122990.
ACCESSION AX105109
VERSION AX105109.1 GI:13921259
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 7 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
misc_feature 1..2
/note="Backbone has phosphorothioate linkages."
misc_feature 3..15
/note="Backbone has phosphodiester linkages."
misc_feature 16..19
/note="Backbone has phosphorothioate linkages."
misc_feature 20
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 2 c 13 g 3 t
ORIGIN
1
//

Query Match      92.0%; Score 18.4; DB 6; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggggacgacgttggtggggg 20
   |||||
Db 1 GGGGGACGATCGTGGGGGG 20

RESULT 4
AX105234
LOCUS AX105234 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 133 from Patent WO0122990.
ACCESSION AX105234
VERSION AX105234.1 GI:13921384
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 133 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
misc_feature 1..2
/note="Backbone has phosphorothioate linkages."
misc_feature 3..13
/note="Backbone has phosphodiester linkages."
misc_feature 14..19
/note="Backbone has phosphorothioate linkages."
misc_feature 20
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN
1
//

Query Match      92.0%; Score 18.4; DB 6; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggggacgacgttggtggggg 20
   |||||
Db 1 GGGGGACGATCGTGGGGGG 20

RESULT 5
AX104805
LOCUS AX104805 21 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 997 from Patent WO0122972.
ACCESSION AX104805
VERSION AX104805.1 GI:13921002
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 997 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:58:05 ; Search time 2778.35 Seconds
(without alignments)
150.640 Million cell updates/sec

Title: US-09-672-126-13
Perfect score: 20
Sequence: 1 gggggacgacgtgtggggg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vl.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htgo.inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

RESULT 1

AX104803 LOCUS AX104803.1
DEFINITION Sequence 995 from Patent WO0122972.
ACCESSION AX104803
VERSION AX104803.1 GI:13921000
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg A.M., Schetter, C. and Vollmer, J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 995 03-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical GmbH (DE)
FEATURES Location/Qualifiers
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 2 a 2 c 13 g 3 t
ORIGIN

ALIGNMENTS

1	20	100.0	20	6	AX104803
2	20	100.0	20	6	AX105115
3	18.4	92.0	20	6	AX105109
4	18.4	92.0	20	6	AX105234
5	18.4	92.0	21	6	AX104805
6	18.4	92.0	21	6	AX104806
7	18.4	92.0	21	6	AX105117
8	18.4	92.0	21	6	AX105118
9	18	90.0	20	6	AX104864
10	17.4	87.0	19	6	AX104767
11	17.4	87.0	20	6	AX104883
12	17.4	87.0	20	6	AX105137
13	16.8	84.0	20	6	AX104804
14	16.8	84.0	20	6	AX104862
15	16.8	84.0	20	6	AX104863
16	16.8	84.0	20	6	AX105116
17	16.8	84.0	20	6	AX105133
18	16.8	84.0	20	6	AX105258
19	16.8	84.0	18655	3	AF125967
20	16.8	84.0	63853	2	AC087663
21	16.8	84.0	76064	2	AC101191
22	16.8	84.0	77376	2	AC094984
23	16.8	84.0	121762	2	AC098672
24	16.8	84.0	129727	9	AC084381
25	16.8	84.0	148183	2	AP004672
26	16.8	84.0	156840	2	AC094846
27	16.8	84.0	242184	2	AC015899
28	16.4	82.0	19	6	AX104879
29	16.4	82.0	19	6	AX105134
30	16.4	82.0	20	6	AX105259
31	16.4	82.0	428	8	AF228471
32	16.4	82.0	1963	8	AF221063
33	16.4	82.0	39640	2	AC014099
34	16.4	82.0	169296	2	AC022198
35	16.4	82.0	174576	9	AC090525
36	16.4	82.0	175775	2	AC026831
37	16.4	82.0	179234	3	AC023682
38	16.4	82.0	187193	2	AC105657
39	16.4	82.0	306135	3	AE003436
40	15.8	79.0	230	8	AY020365
41	15.8	79.0	433	11	AF344013
42	15.8	79.0	1009	33	AC058901
43	15.8	79.0	1500	1	AF255602
44	15.8	79.0	2093	10	RATNGFIC
45	15.8	79.0	4011	10	RATNGFIC3E

linear PAT 30-APR-2001

Coley Pharmaceutical

THIS PAGE BLANK (USPTO)

Qy 1 gggggacgagctcgtcggggg 22
||| ||| ||| ||| ||| |||
Db 35 GGGGGAGGAGCCCCCGGGGAGG 56

Search completed: August 10, 2002, 03:06:06
Job time: 16032 sec

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,751
FILING DATE: 23-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: THOMSON, MARTIN T
REGISTRATION NUMBER: 31432
REFERENCE/DOCKET NUMBER: PPD 50067/US
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3827 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: PROMOTER OF GSTII 27KD SUBUNIT
US-08-718-751-1

Query Match 71.8%; Score 15.8; DB 2; Length 3827;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ggggacgagctcgtcggg 20
|||||
Db 1732 GGGGACGAGCTCGTGGG 1750

RESULT 14
US-09-049-289-6
Sequence 6, Application US/09049289
Patent No. 6066456
GENERAL INFORMATION:
APPLICANT: BRIDGES, IAN G.
APPLICANT: BRIGHT, SIMON W.J.
APPLICANT: GREENLAND, ANDREW J.
APPLICANT: HOLT, DAVID C.
APPLICANT: JEPSON, IAN
APPLICANT: SCHUCH, WOLFGANG W.
TITLE OF INVENTION: PLANT-DERIVED ENZYME AND DNA SEQUENCES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DABRY & CUSHMAN L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/049,289
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/170,294
FILING DATE: 30-DEC-1993
APPLICATION NUMBER: WO PCT/GB92/01187
FILING DATE: 01-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9114259.6
FILING DATE: 02-JUL-1991
ATTORNEY/AGENT INFORMATION:

NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 204218/SEE36438/UST
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3827 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: gst-27 promoter figure 8
US-09-049-289-6

Query Match 71.8%; Score 15.8; DB 3; Length 3827;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ggggacgagctcgtcggg 20
|||||
Db 1732 GGGGACGAGCTCGTGGG 1750

RESULT 15
US-08-253-155A-6
Sequence 6, Application US/08253155A
Patent No. 5691147
GENERAL INFORMATION:
APPLICANT: Gyuris, Jenő
APPLICANT: Draetta, Giulio
TITLE OF INVENTION: CDK4 Binding Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/253,155A
FILING DATE: 02-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-253-155A-6

Query Match 70.9%; Score 15.6; DB 1; Length 252;
Best Local Similarity 81.8%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

; GENERAL INFORMATION:
; APPLICANT: BRIDGES, IAN G.
; APPLICANT: BRIGHT, SIMON W.J.
; APPLICANT: GREENLAND, ANDREW J.
; APPLICANT: HOLT, DAVID C.
; APPLICANT: JEPSON, IAN
; APPLICANT: SCHUCH, WOLFGANG W.
; TITLE OF INVENTION: PLANT-DERIVED ENZYME AND DNA SEQUENCES
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,294
; FILING DATE: 30-DEC-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB92/01187
; FILING DATE: 01-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9114259.6
; FILING DATE: 02-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 204218/SEE36438/UST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3827 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: gst-27 promoter figure 8
US-08-170-294-6

Query Match 71.8%; Score 15.8; DB 1; Length 3827;
Best Local Similarity 89.5%; Pred. NO. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ggggacgagctgcgcggg 20
|||||
Db 1732 GGGGACGAGCTCGCTGGGG 1750

RESULT 12
US-08-664-855-6
; Sequence 6, Application US/08664855
; Patent No. 5866792
; GENERAL INFORMATION:
; APPLICANT: BRIDGES, IAN G
; APPLICANT: BRIGHT, SIMON WJ
; APPLICANT: GREENLAND, ANDREW J
; APPLICANT: HOLT, DAVID C
; APPLICANT: JEPSON, IAN
; APPLICANT: SCHUCH, WOLFGANG W
; TITLE OF INVENTION: PLANT-DERIVED ENZYME AND DNA SEQUENCES,
; TITLE OF INVENTION: AND USES THEREOF

```

```

; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/664,855
; FILING DATE: 17-JUN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/170,294
; FILING DATE: 30-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9114259.6
; FILING DATE: 02-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/01187
; FILING DATE: 01-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 224452/SEE36438USTD1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3827 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: gst-27 promoter figure 8
US-08-664-855-6

Query Match 71.8%; Score 15.8; DB 2; Length 3827;
Best Local Similarity 89.5%; Pred. NO. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ggggacgagctgcgcggg 20
|||||
Db 1732 GGGGACGAGCTCGCTGGGG 1750

RESULT 13
US-08-718-751-1
; Sequence 1, Application US/08718751
; Patent No. 5965387
; GENERAL INFORMATION:
; APPLICANT: JEPSON, IAN
; APPLICANT: GREENLAND, ANDREW J
; APPLICANT: BEVAN, MICHAEL
; APPLICANT: SHEPPARD, HILARY
; TITLE OF INVENTION: A PROMOTER
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA INC.
; STREET: 1200 SOUTH 47TH STREET
; CITY: RICHMOND
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94804-0023

```

```

Query Match      73.6%; Score 16.2; DB 2; Length 3107;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  gggggagcagctcgctggggg 21
      || ||||| ||||| ||||| ||
Db      1643  GGTGGAGCAGCTGGTGGCGG 1663

RESULT      9
US-08-945-056-3
; Sequence 3, Application US/08945056
; Patent No. 6077994
; GENERAL INFORMATION:
; APPLICANT: Coupland, George M.
; APPLICANT: Putterill, Joanna J.
; TITLE OF INVENTION: Genetic control of flowering
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 8th Floor, 1100 No. 6077994th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:

```


APPLICATION NUMBER: 85.178
FILING DATE: 14-AUG-1987
SEQ ID NO:15:
LENGTH: 1040
5492811-15

Query Match 73.6%; Score 16.2; DB 6; Length 1040;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ggggacgagctcgtcgggggg 22
||||| |||||||||
Db 667 GGGGAAGCGCTCGTCGGGTC 647

RESULT 3
US-081-610-1
Sequence 1, Application US/08081610
Patent No. 5445941
GENERAL INFORMATION:
APPLICANT: Yang, Na N
TITLE OF INVENTION: Materials and Methods for Screening
TITLE OF INVENTION: Anti-Osteoporosis or Serum Lipid Lowering Agents
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti and Witcoff, Ltd.
STREET: 10 S. Wacker Dr.
CITY: Chicago
STATE: IL
COUNTRY: U.S.A
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/081,610
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Heaphy, Barbara A
REGISTRATION NUMBER: 34,619
REFERENCE/DOCKET NUMBER: 93,402
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2205 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: misc_RNA
LOCATION: 1..2
OTHER INFORMATION: /note= "Number 1 corresponds to
OTHER INFORMATION: -1362 of TGFB-1 promoter"
FEATURE:
NAME/KEY: misc_RNA
LOCATION: 1363..1365
OTHER INFORMATION: /note= "Corresponds to +1 codon of
OTHER INFORMATION: TGFB-1"
US-081-610-1

Query Match 73.6%; Score 16.2; DB 1; Length 2205;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggggacgagctcgtcgggggg 21

Db 1580 GAGGGACGAGCTGGTCGGGAG 1600

RESULT 4
5168051-1
Patent No. 5168051
APPLICANT: DERYNCK, RIK M.A.; GOEDEL, DAVID V.
TITLE OF INVENTION: NUCLEIC ACID ENCODING TGF-B ITS USES
NUMBER OF SEQUENCES: 21
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/389,929
FILING DATE: 04-AUG-1989
SEQ ID NO:1:
LENGTH: 2537
5168051-1

Query Match 73.6%; Score 16.2; DB 6; Length 2537;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggggacgagctcgtcgggggg 21
||||| |||||||||
Db 219 gggggacgagctcgtcgggggg 239

RESULT 5
PCT-US94-03705-3
Sequence 3, Application PC/TUS9403705
GENERAL INFORMATION:
APPLICANT: Mu-En Lee
APPLICANT: Mark A. Perrella
TITLE OF INVENTION: TRANSFORMING GROWTH
TITLE OF INVENTION: FACTOR- INHIBITS
TITLE OF INVENTION: INDUCIBLE NITRIC OXIDE
TITLE OF INVENTION: SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/03705
FILING DATE: 5 April 1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICANT NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Janis K. Fraser
REGISTRATION NUMBER: Reg. No. 34,819
REFERENCE/DOCKET NUMBER: 05433/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2745
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
PCT-US94-03705-3

REFERENCE

AUTHORS

Sarcocystidae; Neospora.
1 (bases 1 to 441)
Cole, R., Fogarty, S., Tang, K., Howe, D.K., Sibley, L.D., Clifton, S.,
Marra, M., Hillier, L., Pape, D., Martin, J., Wyllie, T., Theising, B.,
Bowers, K., Gibbons, M., Ritter, E., Bennett, J., Ronko, I.,
Tsagareishvili, R., Fedele, M., Belaygorod, L., Franklin, C., Carr, L.M.,
Grow, A., Maguire, L., Wadkins, J., Richey, J., Waterston, R. and
Wilson, R.

TITLE

USDA-WashU Neospora EST Project

JOURNAL

COMMENT

Unpublished (2000)
Contact: Sandy Clifton, Ph.D. - Neospora
USDA-WashU Neospora EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Contact David Sibley (toxoe@borcim.wustl.edu) for further
information relating to organism, libraries, or clone availability.
Seq primer: -40RP from gibco
High quality sequence stop: 317.

FEATURES

source

1..441
Location/Qualifiers
/organism="Neospora caninum"
/strain="NC-1"
/db_xref="taxon:29176"
/clone_lib="NC 1314 Tachyzoite cDNA"
/dev_stage="Tachyzoite"
/lab_host="DH10B (GeneHog, Research Genetics, Inc.)"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; This library was constructed by Steve Fogarty,
Robert Cole, and Keliang Tang at Washington University.
cDNAs were synthesized from poly(A)+ RNA by oligo d(T)
priming, size-selected and directionally cloned into the
Uni-ZAP XR lambda vector (Stratagene). The primary library
was mass excised as phagemids and rescued in SOLR cells.
The plasmid library was recovered from the SOLR cells and
transformed in mass into DH10B (GeneHog, Research Genetics
, Inc.) for sequencing. WARNING: the library may contain a
small percentage of contaminants from human fibroblast
cells."

BASE COUNT 110 a 110 c 129 g 90 t 2 others

ORIGIN

Query Match 80.9%; Score 17.8; DB 10; Length 441;

Best Local Similarity 90.5%; Pred. No. 4.9e+03;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggacgagctcgctcgggg 21

Db 109 GGTGACGAGCTCGTCGGAGG 89

RESULT 15

BE450575/c

LOCUS

BE450575 485 bp mRNA linear EST 18-MAY-2001

EST401462 tomato root, plants pre-anthesis, Cornell University

Lycopersicon esculentum cDNA clone CLEY13P16, mRNA sequence.

ACCESSION

BE450575

VERSION

BE450575.1

GI:9456078

KEYWORDS

EST.

SOURCE

tomato.

Lycopersicon

1 (bases 1 to 485)

van der Hoeven, R.S.,

Upton, J., Hansen, T.,

Fraser, C.M., Martin, G.B.,

Generation of ESTs from tomato root tissue

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

Lycopersicon

1 (bases 1 to 485)

van der Hoeven, R.S., Garvin, D., Matern, A.L., Holt, I.E., Liang, F.,

Upton, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Ronning, C.M.,

Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.

Generation of ESTs from tomato root tissue

JOURNAL

COMMENT

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

source

1..485
Location/Qualifiers
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone_lib="CLEV13P16"
/clone_lib="tomato root, plants pre-anthesis, Cornell
University"
/tissue_type="root"
/dev_stage="plants in pre-anthesis stage"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Tanksley; Tissue supplied by Dave Garvin
(USDA-ARS, Ithaca, NY 14850)."

BASE COUNT 107 a 108 c 115 g 155 t

ORIGIN

Query Match 80.9%; Score 17.8; DB 10; Length 485;

Best Local Similarity 90.5%; Pred. No. 4.9e+03;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ggggacgagctcgctcgggg 22

Db 133 GGGGACGAGCTCGTCGGCGG 113

Search completed: August 10, 2002, 02:11:11

Job time: 13132 sec

XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX Hartmann G, Bratzler RL, Krieg A;
 XX WPI; 2001-290487/30.
 DR Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 XX Claim 201; Page 103; 168pp; English.
 XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX Sequence 24 BP; 3 A; 4 C; 14 G; 3 T; 0 other;
 SQ

Query Match 100.0%; Score 24; DB 22; Length 24;
 Best Local Similarity 100.0%; Pred. No. 0.17;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcgcgtacgtcgcaggggg 24
 |||||
 Db 1 ggggtcgcgtacgtcgcaggggg 24
 |||||

RESULT 2
 AAF99769
 ID AAF99769 standard; DNA; 24 BP.
 XX AAF99769;
 XX 12-JUN-2001 (first entry)
 DE Immunostimulatory nucleic acid #885.
 XX Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX Synthetic.
 OS
 XX WO200122972-A2.
 XX 05-APR-2001.
 XX 25-SEP-2000; 2000WO-US26383.
 XX 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 XX Krieg AM, Schetter C, Vollmer J;
 WPI; 2001-273485/28.
 DR Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 XX Claim 101; Page 57; 338pp; English.

XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious diseases, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 XX Sequence 24 BP; 3 A; 4 C; 14 G; 3 T; 0 other;
 SQ

Query Match 100.0%; Score 24; DB 22; Length 24;
 Best Local Similarity 100.0%; Pred. No. 0.17;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcgcgtacgtcgcaggggg 24
 |||||
 Db 1 ggggtcgcgtacgtcgcaggggg 24
 |||||

RESULT 3
 AAF99838
 ID AAF99838 standard; DNA; 24 BP.
 XX AAF99838;
 XX 12-JUN-2001 (first entry)
 DE Immunostimulatory nucleic acid #954.
 XX Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX Synthetic.
 OS
 XX WO200122972-A2.
 XX 05-APR-2001.
 XX 25-SEP-2000; 2000WO-US26383.
 XX 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 XX Krieg AM, Schetter C, Vollmer J;
 WPI; 2001-273485/28.
 DR Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 XX Claim 101; Page 59; 338pp; English.
 CC The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:21:51 ; Search time 1145.36 Seconds
(without alignments)
35.976 Million cell updates/sec

Title: US-09-672-126-25
Perfect score: 24
Sequence: 1 ggggtcagctacgtcgaggggg 24

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues 3472872

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				N_Geneseq_032802:*			
				1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*	2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*	3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*	4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
				5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*	6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*	7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*	8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
				9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*	10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*	11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*	12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
				13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*	14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*	15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*	16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
				17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*	18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*	19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*	20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
				21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*	22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*	23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*	24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
				SUMMARIES			
Result No.	Score	Query Match	Length DB ID	Description			
1	24	100.0	24 22	AAF98755 Human IFN-alpha im			
2	24	100.0	24 22	AAF99769 Immunostimulatory			
3	24	100.0	24 22	AAF99838 Immunostimulatory			
4	17.6	73.3	1327 20	AA22281 Human secreted pro			
5	16.8	70.0	16766 24	ABL34157 Human immune syste			
6	16.6	69.2	1334 21	AA39409 Rice SYR2 homology			
7	16.6	69.2	1682 15	AAQ73501 DNA encoding Pseud			
8	16.6	69.2	1831 12	AAQ10213 BamHI G-P-J fragme			
9	16.6	69.2	1831 12	AAQ10211 BamHI G-P-J fragme			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	C	10	16.6	69.2	2027	23	AA590594
DNA encoding novel	C	11	16.6	69.2	4689	21	AA887299
S. venezuelae macr	C	12	16.6	69.2	5487	24	ABL33599
Human immune syste	C	13	16.6	69.2	6242	24	ABL34148
Human immune syste	C	14	16.6	69.2	8438	15	AAQ73500
DNA encoding Pseud	C	15	16.6	69.2	13842	21	AA287297
S. venezuelae macr	C	16	16.6	69.2	36778	21	AA887318
S. venezuelae pik	C	17	16.6	69.2	37948	21	AA287285
S. venezuelae pik	C	18	16.6	69.2	38506	21	AA275633
Nucleotide sequenc	C	19	16.6	69.2	38506	21	AA275633
Recombinant cosmid	C	20	16.6	69.2	38506	21	AA275633
Human secreted pro	C	21	16.2	67.5	2081	19	AAV59623
Chemically pretrea	C	22	16.2	67.5	5687	22	AA54316
Total DNA sequenc	C	23	16.2	67.5	30001	18	AA261016
S. aureofaciens DN	C	24	16.2	67.5	30001	20	AA051110
Plant microsatelli	C	25	16.6	66.7	446	21	AA31428
S. commune SC3 cod	C	26	16.6	66.7	679	21	AA95427
Human ORFX ORF947	C	27	16.6	66.7	886	21	AA75392
Actinomyces sp. 3	C	28	16.6	66.7	1059	20	AA72336
Drosophila melanog	C	29	16.6	66.7	1812	23	ABL23633
Nucleotide sequenc	C	30	16.6	66.7	3258	22	AA42269
Drosophila melanog	C	31	16.6	66.7	3812	23	ABL23632
Propionibacterium	C	32	16.6	66.7	6779	23	AA59570
T. versicolor lacc	C	33	16.6	66.7	7986	20	AA23937
Propionibacterium	C	34	15.8	65.8	2170	22	AAH72880
Human cervical can	C	35	15.8	65.8	5224	20	AA232022
Human METH1 relate	C	36	15.8	65.8	5224	22	AA232022
L05390 cDNA clone.	C	37	15.8	65.8	5224	24	AA561263
Human gene regulat	C	38	15.8	65.8	21034	19	AAV62154
HSV-2 strain SB5 C	C	39	15.8	65.8	21034	19	AAV62154
HSV-2 strain SB5 C	C	40	15.8	65.8	26338	19	AAV62134
HSV-2 strain SB5 C	C	41	15.8	65.8	26338	19	AAV62134
HSV-2 strain SB5 C	C	42	15.8	65.8	117213	19	AAV62176
HSV-2 strain SB5 C	C	43	15.8	65.8	117213	19	AAV62176
Human herpesvirus	C	44	15.8	65.8	154746	24	AA25519
Human herpesvirus	C	45	15.8	65.8	154746	24	AA25519

ALIGNMENTS

RESULT	1
AAF98755	ID
AAF98755	standard; DNA; 24 BP.
XX	AC
XX	AAF98755;
XX	11-JUN-2001 (first entry)
XX	Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 25.
DE	Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
DE	virial infection; phosphorothioate backbone; palindrome; cancer; ds.
KW	Synthetic.
XX	OS
XX	Key
XX	modified_base
XX	Location/Qualifiers
XX	1..2
XX	/tag= a
XX	/mod_base= "OTHER"
XX	/note= "phosphorothioate linkage"
XX	19..23
XX	modified_base
XX	/tag= b
XX	/mod_base= "OTHER"
XX	/note= "phosphorothioate linkage"
XX	WO200122990-A2.
XX	05-APR-2001.
XX	27-SEP-2000; 2000WO-US26527.
XX	27-SEP-1999; 99US-0156147.

THIS PAGE BLANK (UBPTO)

```

* 63547 63646: gap of unknown length
* 63647 64725: contig of 1079 bp in length
* 64726 64825: gap of unknown length
* 64826 66053: contig of 1228 bp in length
* 66054 66153: gap of unknown length
* 66154 67199: contig of 1046 bp in length
* 67200 67299: gap of unknown length
* 67300 68495: contig of 1196 bp in length
* 68496 68595: gap of unknown length
* 68596 69703: contig of 1108 bp in length
* 69704 69803: gap of unknown length
* 69804 71263: contig of 1460 bp in length
* 71264 71363: gap of unknown length
* 71364 72513: contig of 1150 bp in length
* 72514 72613: gap of unknown length
* 72614 73916: contig of 1303 bp in length
* 73917 74016: gap of unknown length
* 74017 75254: contig of 1238 bp in length
* 75255 75354: gap of unknown length
* 75355 76397: contig of 1043 bp in length
* 76398 76497: gap of unknown length
* 76498 77506: contig of 1009 bp in length.

```

FEATURES

```

Location/Qualifiers
1..77506
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-39G7"

```

BASE COUNT 21465 a 16104 c 15986 g 19612 t 4339 others

ORIGIN

```

Query Match 73.3% Score 17.6; DB 2; Length 77506;
Best Local Similarity 83.3%; Pred. No. 7.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcgcgtacgtcgagggggg 24
    ||||| ||| ||||| |||||
Db 61973 GGGGCGAGGTACGTGTTGGGGG 61950

```

Search completed: August 10, 2002, 02:58:29
Job time: 15715 sec

KEYWORDS

HTG; HTGS_PHASE1.

Norway rat.

SOURCE

Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 77506)

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Alsbrooks, S.L., Anaratunge, H.C., Are, J.R., Banks, T., Barbacia, J.,
Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Delgado, O.,
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C.,
Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, F., Johnson, R., Jolivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Lowsegh, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
Martinez, E., Massey, M., Mawhney, E., McLeod, M.P., Meador, M.,
Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S.,
Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M.,
Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshtari, N.,
Slisdon, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Thomas, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Watling, S., Williams, G., Williams, A., Wleczky, R., Wooden, S.,
Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission

2 (bases 1 to 77506)

Worley, K.C.

Direct Submission

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15799424.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GESR

Center clone name: CH230-3967

----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to

findPhrapList

Consensus quality: 54532 bases at least Q40

Consensus quality: 60036 bases at least Q30

Consensus quality: 64148 bases at least Q20

Estimated insert size: 44611; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-gel estimation

Quality coverage: 0.5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 43 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 5100: contig of 5100 bp in length
* 5101 5200: gap of unknown length
* 7876: contig of 2676 bp in length
* 7877 7976: gap of unknown length
* 7977 10385: contig of 2409 bp in length
* 10386 10485: gap of unknown length
* 10486 11878: contig of 1393 bp in length
* 11879 11978: gap of unknown length
* 11979 14632: contig of 2654 bp in length
* 14633 14732: gap of unknown length
* 14733 17115: contig of 2983 bp in length
* 17116 17715: gap of unknown length
* 17716 20279: contig of 2464 bp in length
* 20280 20379: gap of unknown length
* 20380 22926: contig of 2547 bp in length
* 22927 23026: gap of unknown length
* 23027 25051: contig of 2024 bp in length
* 25051 25150: gap of unknown length
* 25151 26414: contig of 1264 bp in length
* 26415 26514: gap of unknown length
* 26515 28077: contig of 1563 bp in length
* 28078 28177: gap of unknown length
* 28178 30789: contig of 2612 bp in length
* 30790 30889: gap of unknown length
* 30890 33184: contig of 2295 bp in length
* 33185 33284: gap of unknown length
* 33285 35453: contig of 2169 bp in length
* 35454 35553: gap of unknown length
* 35554 37192: contig of 1639 bp in length
* 37193 37292: gap of unknown length
* 37293 39022: contig of 1730 bp in length
* 39023 39122: gap of unknown length
* 39123 40624: contig of 1502 bp in length
* 40625 40724: gap of unknown length
* 40725 42312: contig of 1588 bp in length
* 42313 42412: gap of unknown length
* 42413 43968: contig of 1556 bp in length
* 43969 44068: gap of unknown length
* 44069 45192: contig of 1124 bp in length
* 45193 45292: gap of unknown length
* 45293 47370: contig of 2078 bp in length
* 47371 47470: gap of unknown length
* 47471 49119: contig of 1649 bp in length
* 49120 49219: gap of unknown length
* 49220 51008: contig of 1789 bp in length
* 51009 51108: gap of unknown length
* 51109 52191: contig of 1083 bp in length
* 52192 52976: contig of 1185 bp in length
* 52977 53477: gap of unknown length
* 53478 53576: contig of 1580 bp in length
* 53577 55256: gap of unknown length
* 55257 56654: contig of 1398 bp in length
* 56655 56754: gap of unknown length
* 56755 58031: contig of 1277 bp in length
* 58032 58131: gap of unknown length
* 58132 59617: contig of 1486 bp in length
* 59618 59718: gap of unknown length
* 59719 60843: contig of 1126 bp in length
* 60844 60944: gap of unknown length
* 60945 61994: contig of 1051 bp in length
* 61995 62094: gap of unknown length
* 62095 63546: contig of 1452 bp in length

```
* 4110 4209: gap of 100 bp
* 4210 4999: contig of 790 bp in length
* 5000 5099: gap of 100 bp
* 5100 5841: contig of 742 bp in length
* 5842 5941: gap of 100 bp
* 5942 6645: contig of 704 bp in length
* 6646 6745: gap of 100 bp
* 6746 7470: contig of 725 bp in length
* 7471 7570: gap of 100 bp
* 7571 8326: contig of 756 bp in length
* 8327 8426: gap of 100 bp
* 8427 9156: contig of 730 bp in length
* 9157 9256: gap of 100 bp
* 9257 9978: contig of 722 bp in length
* 9979 10078: gap of 100 bp
* 10079 10842: contig of 764 bp in length
* 10843 10942: gap of 100 bp
* 10943 11702: contig of 760 bp in length
* 11703 11802: gap of 100 bp
* 11803 12561: contig of 759 bp in length
* 12562 12661: gap of 100 bp
* 12662 13471: contig of 810 bp in length
* 13472 13571: gap of 100 bp
* 13572 14348: contig of 777 bp in length
* 14349 14448: gap of 100 bp
* 14449 15162: contig of 714 bp in length
* 15163 15262: gap of 100 bp
* 15263 15973: contig of 711 bp in length
* 15974 16073: gap of 100 bp
* 16074 16804: contig of 731 bp in length
* 16805 16904: gap of 100 bp
* 16905 17699: contig of 795 bp in length
* 17700 17799: gap of 100 bp
* 17800 18554: contig of 755 bp in length
* 18555 18654: gap of 100 bp
* 18655 19368: contig of 714 bp in length
* 19369 19468: gap of 100 bp
* 19469 20294: contig of 826 bp in length
* 20295 20394: gap of 100 bp
* 20395 21179: contig of 785 bp in length
* 21180 21279: gap of 100 bp
* 21280 22023: contig of 744 bp in length
* 22024 22123: gap of 100 bp
* 22124 22845: contig of 722 bp in length
* 22846 22945: gap of 100 bp
* 22946 23656: contig of 711 bp in length
* 23657 23756: gap of 100 bp
* 23757 24486: contig of 730 bp in length
* 24487 24586: gap of 100 bp
* 24587 25314: contig of 728 bp in length
* 25315 25414: gap of 100 bp
* 25415 26117: contig of 703 bp in length
* 26118 26217: gap of 100 bp
* 26218 26957: contig of 740 bp in length
* 26958 27057: gap of 100 bp
* 27058 27765: contig of 708 bp in length
* 27766 27865: gap of 100 bp
* 27866 28595: contig of 730 bp in length
* 28596 28695: gap of 100 bp
* 28696 29365: contig of 670 bp in length
* 29366 29465: gap of 100 bp
* 29466 30265: contig of 800 bp in length
* 30266 30365: gap of 100 bp
* 30366 31074: contig of 709 bp in length
* 31075 31174: gap of 100 bp
* 31175 31898: contig of 724 bp in length
* 31899 31998: gap of 100 bp
* 31999 32763: contig of 765 bp in length
* 32764 32863: gap of 100 bp
* 32864 33598: contig of 735 bp in length
* 33599 33698: gap of 100 bp
* 33699 34424: contig of 726 bp in length
* 34425 34524: gap of 100 bp
```

```
* 34525 35253: contig of 729 bp in length
* 35254 35353: gap of 100 bp
* 35354 36104: contig of 751 bp in length
* 36105 36204: gap of 100 bp
* 36205 36954: contig of 750 bp in length
* 36955 37054: gap of 100 bp
* 37055 37775: contig of 721 bp in length
* 37776 37875: gap of 100 bp
* 37876 38615: contig of 740 bp in length
* 38616 38715: gap of 100 bp
* 38716 39448: contig of 733 bp in length
* 39449 39548: gap of 100 bp
* 39549 40303: contig of 755 bp in length
* 40304 40403: gap of 100 bp
* 40404 41141: contig of 738 bp in length
* 41142 41241: gap of 100 bp
* 41242 41945: contig of 704 bp in length
* 41946 42045: gap of 100 bp
* 42046 42765: contig of 720 bp in length
* 42766 42865: gap of 100 bp
* 42866 43623: contig of 758 bp in length
* 43624 43723: gap of 100 bp
* 43724 44430: contig of 707 bp in length
* 44431 44530: gap of 100 bp
* 44531 45247: contig of 717 bp in length
* 45248 45347: gap of 100 bp
* 45348 46064: contig of 717 bp in length
* 46065 46164: gap of 100 bp
* 46165 46910: contig of 746 bp in length
* 46911 47010: gap of 100 bp
* 47011 47730: contig of 720 bp in length
* 47731 47830: gap of 100 bp
* 47831 48548: contig of 718 bp in length
* 48549 48648: gap of 100 bp
* 48649 49414: contig of 766 bp in length
* 49415 49514: gap of 100 bp
* 49515 50251: contig of 737 bp in length
* 50252 50351: gap of 100 bp
* 50352 51083: contig of 732 bp in length
* 51084 51183: gap of 100 bp
* 51184 51909: contig of 726 bp in length
* 51910 52009: gap of 100 bp
* 52010 52746: contig of 737 bp in length
* 52747 52846: gap of 100 bp
* 52847 53579: contig of 733 bp in length
* 53580 53679: gap of 100 bp
* 53680 54405: contig of 726 bp in length
* 54406 54505: gap of 100 bp
* 54506 55248: contig of 743 bp in length
* 55249 55348: gap of 100 bp
* 55349 56088: contig of 740 bp in length
* 56089 56188: gap of 100 bp
* 56189 56926: contig of 738 bp in length
* 56927 57026: gap of 100 bp
* 57027 57749: contig of 723 bp in length
* 57750 57849: gap of 100 bp
```

Query Match

73.3%; Score 17.6; DB 2; Length 65172;

Best Local Similarity 83.3%; Pred. NO. 7.4e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 ggggtcgcgtacgtcgcagggggg 24

||||| | ||||| |||||

Db 13355 GGGGTGTCGTACGTCTCGGGGGG 13332

RESULT 15

AC096238/c

LOCUS

DEFINITION

AC096238

AC096238

AC096238.3

AC096238 Rattus norvegicus clone CH230-3967, *** SEQUENCING IN PROGRESS ***,
43 unordered pieces.
AC096238
AC096238.3 GI:17943929

77506 bp DNA linear HTG 20-DEC-2001

Matches		20; Conservative	0; Mismatches	4; Indels	0; Gaps	0;
QY	1	ggggtcagctacgtcaggggg 24 				
Db	13022	GGAGGCGCGTACGTGAGCGGG 13045				
RESULT	13					
LOCUS	SCK7	36539 bp DNA linear BCT 24-JAN-2001				
DEFINITION		Streptomyces coelicolor cosmid K7.				
ACCESSION		AL391754				
VERSION		AL391754.1 GI:9967654				
KEYWORDS		ABC transporter ATP-binding protein; ahpc; ahpd; amino acid permease; ATP-binding protein; dehydrogenase; exoribonuclease large subunit; exoribonuclease small subunit; fumb; fumarate hydratase class I; fumb; fumarate hydratase C; glycosyltransferase; hydrolase; integral membrane protein; membrane protein; oxyR; penicillin-binding protein; secreted protein; wbli.				
SOURCE		Streptomyces coelicolor.				
ORGANISM		Streptomyces coelicolor Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomyceinae; Streptomycetaceae; Streptomyces.				
REFERENCE		1 (bases 1 to 36539)				
AUTHORS		Redenbach,M., Kieser,H.M., Denapaiter,D., Eichner,A., Cullum,J., Kinashi,H. and Hopwood,D.A.				
TITLE		A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome				
JOURNAL		Mol. Microbiol. 21 (1), 77-96 (1996)				
MEDLINE		97000351				
REFERENCE		2 (bases 1 to 36539)				
AUTHORS		Seeger,K.J. and Harris,D.				
JOURNAL		Unpublished				
REFERENCE		3 (bases 1 to 36539)				
AUTHORS		Cerdeno,A.M., Parkhill,J., Barrell,B.G. and Rajandream,M.A.				
TITLE		Direct Submission				
JOURNAL		Submitted (31-AUG-2000) Streptomyces coelicolor sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK				
COMMENT		Notes: Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC and Beowulf Genomics Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web. (URL; http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are numbered using the following system eg SC7B7.01c, SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary strand). The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS. Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at http://www.nih.go.jp/jun/cgi-bin/frameplot.pl . CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or att) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon. IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid K7 overlaps cosmid K15 on the AseI-K genomic restriction fragment.				
FEATURES		source	Location/Qualifiers			
		source	1..36539 /organism="Streptomyces coelicolor" /db_xref="taxon:1902" 1..36539 /organism="Streptomyces coelicolor A3(2)" /strain="A3(2)" /db_xref="taxon:100228" /clone="cosmid K7" 1..102 /note="nominal overlap with Streptomyces coelicolor cosmid SK15" 118..316 /note="high G+C content region (81.4%)" 894..2219 /gene="SCK7.01" 894..2219 /gene="SCK7.01" /note="SCK7.01, possible ATP-binding protein, len: 441 aa; similar to TR:Q9X792 (EMBL:AL049491) Mycobacterium leprae putative ATP-binding protein MLCB1222.20, 433 aa; fasta scores: opt: 2024 z-score: 2236.8 E(): 0; 72.2% identity in 428 aa overlap. Contains and match to prosite entry PS00017 ATP/GTP-binding site motif A (P-loop)" /codon_start=1 /transl_table=11 /product="putative ATP-binding protein" /protein_id="CAC05873.1" /db_xref="GI:9967655" /translation="MVTSTKRHKPDRTYVLDTSVLLADPNALNRFDEHEVVLPIVV TELEAKRHHPGLGYFARQALRLLEDFVRHGRDLDAIPIDGLGGTVRVELNHSDFSVL PTGYRLGNDNSRLFAVARNLOAEGDVTWVKDPLRIKASSVGLLAERYRAELATD ASGWTGMSLELTLPGEQDVLFEEGRVYVPEAAGLPVHTGLTIQSERGALGRVTPDGN VRLVGRDREAFGICKRSAEQRIALLDLLDPVGIYSGMGRAGTGSALACGLLEAVL ERRQKQVVFRLPFAVGQELGYPGSEAKMSPQAQVFTLSAVTSREVIEWEVA RMLVEVPLTHIRGKSLHDAFVIVDEAQSLERNVLLTVLSRIGANSRVVLTLDHVAQRD NLRGVRYGVVAVVEKLGHPFLFAHTVLTTRRSRSQIAALVTEMLEDGH" 1662..1685 /gene="SCK7.01" /note="PS00017 ATP/GTP-binding site motif A (P-loop)" 2649..3365 /gene="SCK7.02" 2649..3365 /gene="SCK7.02" /note="SCK7.02, possible secreted protein, len: 238 aa; similar to C-terminal region of TR:CA933446 (EMBL:AL357591) Streptomyces coelicolor hypothetical 16.0 kDa protein SCC53.17, 154 aa; fasta scores: opt: 477 z-score: 484.5 E(): 1.6e-19; 64.4% identity in 101 aa overlap. Contains possible N-terminal region signal peptide sequence and possible coiled-coil region at aprox. residues 82..123" /codon_start=1 /transl_table=11 /product="putative secreted protein" /protein_id="CAC05874.1" /db_xref="GI:9967656" /translation="MLEGNRVSRISVRGFAVASATATVAVGVSAGSVAGVNNDDA EATAGTLLADIPMGDAQVQTASLTQADVQAIADAARAKKAEAEARAKAAATAV DKQKAKAAQAAQRAQAAAKRAKRETSFAVQSSYSTSQIAQARQWPGQF QCFSNIVNHSSWNYQAVNASGAYGLFQALPAGKYASAGADWRTNPATQIKWGLSYM DNRYGSPCDANAFWQAHWY" 3517..3520 3528..4925 /gene="SCK7.03" 3528..4925 /gene="SCK7.03" /note="SCK7.03, possible integral membrane protein, len: 465 aa; similar to SN:Y205_MYCTU (EMBL:AL021928) Mycobacterium tuberculosis hypothetical 38.0 kDa protein MTU033.13, 367 aa; fasta scores: opt: 892 z-score: 833.4 E(): 0; 41.1% identity in 343 aa overlap. Contains Pfam match to entry PF01594 UPF0118, Domain of unknown function			
		misc_feature				
		misc_feature				
		gene				
		CDS				
		misc_feature				
		gene				
		CDS				
		RBS				
		gene				
		CDS				

upstream initiation codon.
 IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid E41.

```

FEATURES
  source
    1..36028
      /organism="Streptomyces coelicolor"
      /db_xref="taxon:1902"
      complement(1..951)
      /gene="SCE41.01c"
      complement(<1..951)
      /gene="SCE41.01c"
      /note="SCE41.01c, probable oxidoreductase (fragment), len:
      >317 aa; similar to SW:DHNA_ECOLI (EMBL:V00306)
      Escherichia coli NADH dehydrogenase (EC 1.6.99.3) Ndh, 433
      aa; fasta scores: opt: 361 z-score: 417.8 E(): 9.1e-16;
      27.6% identity in 330 aa overlap and to TR:CAB92372
      (EMBL:AL356612) Streptomyces coelicolor putative NADH
      dehydrogenase SCD72A.05, 442 aa; fasta scores: opt: 1185
      z-score: 1267.3 E(): 0; 57.0% identity in 321 aa overlap.
      Contains Pfam match to entry PF00070 pyr_redox, Pyridine
      nucleotide-disulphide oxidoreductase"
      /codon_start=1
      /transl_table=11
      /product="putative oxidoreductase (fragment)"
      /protein_id="CAC09533.1"
      /db_xref="GI:10241775"
      /translation="MRYGEATVVDPRSYMYQPLPEAAAGSISPRHVVPLRVL
      PRAEVLGRTTIDQKRVATVAPLVEGEAYELPFDFLYIAMGAVSFTPIPLAEQGI
      GMKIEESLGNLEQIDKADSTDEIRRKALTFVFGGFGAGAEITGEVDNAR
      DAAKYNNYREDMRFILVDAADKILPEVGRKQYGRKHELEGRGVYVLTSDSCV
      DGHVVLKNGLEVDSTIVMTAGVKPNPALAREGLPLGPRGHVDTQATLQVQGTDTIWA
      AGDNAQVPLVGRKAGNENACPPNAQHALRQAKVLGDNVI"
      1..101
      /note="nominal overlap with Streptomyces coelicolor cosmid
      SCE25"

  source
    1..36028
      /organism="Streptomyces coelicolor A3(2)"
      /strain="A3(2)"
      /db_xref="taxon:100226"
      /clone="cosmid E41"
      complement(112..798)
      /gene="SCE41.01c"
      /note="Pfam match to entry PF00070 pyr_redox, Pyridine
      nucleotide-disulphide oxidoreductase, score 76.80, E-value
      6.7e-22"
      complement(1477..2418)
      /gene="SCE41.02c"
      complement(1477..2418)
      /gene="SCE41.02c"
      /note="SCE41.02c, possible hydrolase, len: 313 aa; similar
      to N-terminal region of SW:GPPA_ECOLI (EMBL:M87049)
      Escherichia coli guanosine-5'-triphosphate,3'-diphosphate
      pyrophosphatase (EC 3.6.1.40) GppA, 494 aa; fasta scores:
      opt: 311 z-score: 357.2 E(): 2.2e-12; 28.4% identity in
      310 aa overlap"
      /codon_start=1
      /transl_table=11
      /product="putative hydrolase"
      /protein_id="CAC09534.1"
      /db_xref="GI:10241776"
      /translation="MTRVAADCNTSIRLVADADPATGELTDLDRRMTIVRLGGGV
      DRTGLPALERTAEACREYAEVVKHGAERLRFVATSGASDAENRDPDFVRGLDIL
      GVEPEVIGDQAEFSFTGAKELTGKADLPYLVDIGGGSTFEVVGEDHVRARS
      VDVGVVTERHLVRDGAVTDPPTAEQVNAEMADIEALDLAGRVVPGIEATLVGLA
      GSVTVYSAIAQELPEYDSAAIHHSRVRDRREITDWLLASTHAERAAVASHMPGRVD
      VTAAGSLVLAIMERTGAEVHVSEHILDGIAWSIA"
      complement(2415..2975)
      /gene="SCE41.03c"
      complement(2415..2975)
      /gene="SCE41.03c"
      /note="SCE41.03c, conserved hypothetical protein, len: 186

  misc_feature
    aa; similar to TR:P96375 (EMBL:Z92539) Mycobacterium
    tuberculosis hypothetical 16.6 kDa protein MTCY10G2.24c,
    155 aa; fasta scores: opt: 632 z-score: 730.7 E(): 0;
    67.6% identity in 136 aa overlap"
    /codon_start=1
    /transl_table=11
    /product="conserved hypothetical protein"
    /protein_id="CAC09535.1"
    /db_xref="GI:10241777"
    /translation="MQTPPTPTTPTPTPTDADVAAFKQQLGRPPRGLRAIAHRCPCGQ
    DVVETAPRLPDGTPFTLYLTCPKAASAIQLEANGVMKEMTERLATDPELAAYRA
    AHEDYIRRRDEIEELTFPSAGGMDPRVKLHVLAHSLAAGPGVNPGLGEATAMLPE
    WNRKGPCVPTFEQTDDETGTQEDAQ"
    complement(2425..2434)
    /gene="SCE41.03c"
    complement(3061..3585)
    /gene="SCE41.04c"
    complement(3061..3585)
    /gene="SCE41.04c"
    /note="SCE41.04c, hypothetical protein, len: 174 aa;
    similar to TR:P96376 (EMBL:Z92539) Mycobacterium
    tuberculosis hypothetical 24.6 kDa protein MTC10G2.25c,
    228 aa; fasta scores: opt: 273 z-score: 332.6 E():
    5.1e-11; 35.6% identity in 149 aa overlap. Contains
    possible coiled-coil region at approx residues 87..106"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="CAC09536.1"
    /db_xref="GI:10241778"
    /translation="MCGDVRACTGGGDOMAVKDRDRFSTATIRIIGEQTAARVYRSQ
    TRQRARSRLTGRAALLAMVLCSLVVALAYPIQYVAQRAEIALDLQRETRQRVED
    LRLKARWDDAYAEQQVRLRLHYVMPGETGFVVDPPEAAEQFRAGAADRWPYQNV
    WGVKADAVARRO"
    complement(3639..4919)
    /gene="eno"
    /note="SCE41.05c"
    complement(3639..4919)
    /gene="eno"
    /note="SCE41.05c, eno, enolase, len: 426 aa; similar to
    SW:ENO_ECOLI (EMBL:X82400) Escherichia coli enolase (EC
    4.2.1.11) Eno, 431 aa; fasta scores: opt: 1603 z-score:
    1807.6 E(): 0; 60.6% identity in 419 aa overlap. Contains
    Pfam match to entry PF00113 enolase, Enol-ase and match to
    Prosite entry PS00164 Enolase signature"
    /codon_start=1
    /transl_table=11
    /product="enolase"
    /protein_id="CAC09537.1"
    /db_xref="GI:10241779"
    /translation="MPSIDVWVAREILDSRGNPTVEVGLDGGSTGRAVPSGASTG
    AFEATLRDGPDSRYLKGKGVKAVLAVIQIGPELVGYDATEQRLIDQAMFDLADTN
    KESLGANALYGLSLAVAHAAASASDLPLFRLGGPNAHLPLVPMNLLNGSHADSNV
    DLOEFMIAPIGAESSEALRWGAEVTHTLKVLKNGLATGLGDEGFPAPNGSREA
    LLILEIAKEAGTYGEEQIALADVAASEFYKDGSAFEGKKNRSAAEYEAELVEA
    YPLVLEPLDFEDDGMWNTITAKLDGVLQVLDGLFVNPERLARGIEENSANALLY
    KYNQIGSLTETILDVAELQNGFKCMNSHRSGETEDVTIADLAVATNCQIKTGAPAR
    SERVAKYNOLLRIEIEILDAAVYAGRSAPFPFKG"
    complement(3648..4916)
    /gene="eno"
    /note="Pfam match to entry PF00113 enolase, Enol-ase,
    score 835.40, E-value 1.9e-247"
    complement(3885..3926)
    /gene="eno"
    /note="PS00164 Enolase signature"
    complement(4926..4931)
    /gene="eno"
    complement(5181..5915)
    /gene="SCE41.06c"
    complement(5181..5915)
    /gene="SCE41.06c"

  RBS
  gene
  CDS

  misc_feature
    Query Match 73.3%; Score 17.6; DB 1; Length 36028;
    Best Local Similarity 83.3%; Pred. No. 8e+02;

```

* 76906 77005: gap of unknown length
* 77006 80729: contig of 3724 bp in length
* 80730 80829: gap of unknown length
* 80830 83036: contig of 2207 bp in length
* 83037 83136: gap of unknown length
* 83137 85274: contig of 2138 bp in length
* 85275 85374: gap of unknown length
* 85375 87391: contig of 2017 bp in length
* 87392 87491: gap of unknown length
* 87492 89261: contig of 1770 bp in length
* 89262 89361: gap of unknown length
* 89362 91714: contig of 2353 bp in length
* 91715 91814: gap of unknown length
* 91815 93411: contig of 1597 bp in length
* 93412 93511: gap of unknown length
* 93512 94553: contig of 1142 bp in length
* 94554 94753: gap of unknown length
* 94754 96569: contig of 1816 bp in length
* 96570 96669: gap of unknown length
* 96670 99223: contig of 2554 bp in length
* 99224 99323: gap of unknown length
* 99324 101269: contig of 1946 bp in length
* 101270 101369: gap of unknown length
* 101370 102632: contig of 1263 bp in length
* 102633 102732: gap of unknown length
* 102733 105679: contig of 2947 bp in length
* 105680 105779: gap of unknown length
* 105780 107242: contig of 1453 bp in length
* 107243 107342: gap of unknown length
* 107343 109240: contig of 1898 bp in length
* 109241 109340: gap of unknown length
* 109341 110867: contig of 1527 bp in length
* 110868 110967: gap of unknown length
* 110968 112436: contig of 1459 bp in length
* 112437 112536: gap of unknown length
* 112537 114392: contig of 1856 bp in length
* 114393 114492: gap of unknown length
* 114493 115832: contig of 1340 bp in length
* 115833 115932: gap of unknown length
* 115933 117539: contig of 1607 bp in length
* 117540 117639: gap of unknown length
* 117640 118859: contig of 1220 bp in length
* 118860 118959: gap of unknown length
* 118960 120294: contig of 1335 bp in length
* 120295 120394: gap of unknown length
* 120395 122349: contig of 1955 bp in length
* 122350 122449: gap of unknown length
* 122450 123744: contig of 1295 bp in length
* 123745 123844: gap of unknown length
* 123845 125203: contig of 1359 bp in length
* 125204 125303: gap of unknown length
* 125304 126762: contig of 1459 bp in length
* 126763 126862: gap of unknown length
* 126863 128229: contig of 1367 bp in length
* 128230 128329: gap of unknown length
* 128330 129763: contig of 1434 bp in length
* 129764 129863: gap of unknown length
* 129864 131418: contig of 1555 bp in length
* 131419 131518: gap of unknown length
* 131519 132954: contig of 1436 bp in length
* 132955 133054: gap of unknown length
* 133055 134648: contig of 1594 bp in length
* 134649 134748: gap of unknown length
* 134749 136417: contig of 1669 bp in length
* 136418 136517: gap of unknown length
* 136518 137610: contig of 1093 bp in length
* 137611 137710: gap of unknown length
* 137711 139054: contig of 1344 bp in length
* 139055 139154: gap of unknown length
* 139155 140910: contig of 1756 bp in length
* 140911 141010: gap of unknown length
* 141011 142708: contig of 1698 bp in length
* 142709 142709: gap of unknown length

* 142809 143935: contig of 1127 bp in length
* 143936 144035: gap of unknown length
* 144036 145373: contig of 1538 bp in length
* 145374 145673: gap of unknown length

Query Match 74.2%; Score 17.8; DB 2; Length 158580;
Best Local Similarity 90.5%; Pred. No. 5.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 gtcgacgtacgtcgagggggg 24
||||| ||||||| |||||
Db 117398 GTCGCGTACGTCGCGGGG 117418

RESULT 12
SCE41
LOCUS Streptomyces coelicolor cosmid E41. 36028 bp DNA linear BCT 02-NOV-2000
DEFINITION Streptomyces coelicolor cosmid E41.
ACCESSION AL442120
VERSION AL442120.1 GI:10241774
KEYWORDS ABC transport system ATP-binding protein; ABC transport system integral membrane protein; cytochrome P450 hydroxylase; eno, enolase; hydrolase; integral membrane protein; lipoprotein; nucleotidyltransferase; oxidoreductase; pkab; secreted protein; sensor kinase; transcriptional-repair coupling factor.
SOURCE Streptomyces coelicolor.
ORGANISM Streptomyces coelicolor.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
1 (bases 1 to 36028)
AUTHORS Kadenbach, M., Kieser, H.M., Denapate, D., Eichner, A., Cullum, J., Kinashi, H. and Hopwood, D.A.
TITLE A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome
JOURNAL Mol. Microbiol. 21 (1), 77-96 (1996)
MEDLINE 97000351
REFERENCE 2 (bases 1 to 36028)
AUTHORS Saunders, D.C. and Harris, D.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 36028)
AUTHORS Cerdeno, A.M., Parkhill, J., Barrell, B.G. and Rajandream, M.A.
TITLE Direct Submission
JOURNAL Submitted (19-SEP-2000) Streptomyces coelicolor sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK

Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web
(URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are numbered using the following system eg SC7B7.01c. SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary strand).
The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov.jp/jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or att) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most

```

/misc_feature
/db_xref="GI:9843819"
/translation="MSEQQNLRRRRSETRRLVQAHVRLFTDHCYDAVTADVAEA
AGVSAMTVYRHFPTKEDLVLPQALIAEHVAASAAQPLVRRVGSALIDATNWTG
GNGDEQAANERFLDCLRLVMTSTPALRHLDSQYALQOAIIVDALGDGDAFRA
QAATSCIAAMHTALTRWYDDGHTKLPDLIARALTAFSGDDAVAFRRKG"
complement(3547..3666)
/gene="SC8E7, 06c"
/notes="Pfam match to entry PF00440 tetr, Bacterial
regulatory proteins, tetr family, score 57.90, E-value
2e-13"

Query Match      74.2%; Score 17.8; DB 1; Length 39741;
Best Local Similarity 90.5%; Pred. No. 6.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ggtcgacgtacgtcgaggggg 23
    ||| | ||||| ||||| |||
Db 30250 GGTGGCGGTACGTGCGAGGGG 30270

RESULT 11
AC095115
LOCUS
DEFINITION
Rattus norvegicus clone CH230-7L10, *** SEQUENCING IN PROGRESS ***,
65 unordered pieces.
AC095115
AC095115.2 GI:17941992
HTG: HTGS_PHASE1.
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 158580)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brivana,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Deigado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louisghe,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
Ogulu,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,

```

```

Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 158580)
Worley,K.C.
Direct Submission
Submitted (16-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15625669.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GCLV
Center clone name: CH230-7L10
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 113267 bases at least Q40
Consensus quality: 126964 bases at least Q30
Consensus quality: 137003 bases at least Q20
Estimated insert size: 126268; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.1x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 65 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
* 1 7995: contig of 7995 bp in length
* 7996 8095: gap of unknown length
* 8096 15006: contig of 6911 bp in length
* 15007 15106: gap of unknown length
* 15107 21395: contig of 6289 bp in length
* 21396 21495: gap of unknown length
* 21496 28154: contig of 6659 bp in length
* 28155 28254: gap of unknown length
* 28255 34142: contig of 5888 bp in length
* 34143 34242: gap of unknown length
* 34243 40629: contig of 6387 bp in length
* 40630 40729: gap of unknown length
* 40730 44379: contig of 3650 bp in length
* 44380 44479: gap of unknown length
* 44480 47531: contig of 3052 bp in length
* 47532 47631: gap of unknown length
* 47632 51773: contig of 4142 bp in length
* 51773 51873: gap of unknown length
* 51874 54236: contig of 2363 bp in length
* 54237 54336: gap of unknown length
* 54337 57388: contig of 3052 bp in length
* 57389 57488: gap of unknown length
* 57489 60016: contig of 2528 bp in length
* 60017 60116: gap of unknown length
* 60117 63398: contig of 3282 bp in length
* 63399 63498: gap of unknown length
* 63499 66188: contig of 2690 bp in length
* 66189 66288: gap of unknown length
* 66289 68969: contig of 2681 bp in length
* 68970 69069: gap of unknown length
* 69070 72325: contig of 3256 bp in length
* 72326 72426: gap of unknown length
* 72427 74810: contig of 2385 bp in length
* 74811 74910: gap of unknown length
* 74911 76905: contig of 1995 bp in length

```

CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov/ip/jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or att) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon. IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 8E7 lies towards the end of the chromosome and overlaps cosmid 10B8A.

```

FEATURES             Location/Qualifiers
     source            1..39741
                        /organism="Streptomyces coelicolor A3(2)"
                        /strain="A3(2)"
                        /db_xref="taxon:100226"
                        /clone="cosmid 8E7"
     gene              1..181
                        /gene="SC8E7.01"
                        <1..181
                        /gene="SC8E7.01"
                        /note="SC8E7.01, possible membrane protein, partial CDS,
                        len: 259 aa. Contains possible membrane spanning
                        hydrophobic domains."
     CDS                /codon_start=2
                        /transl_table=11
                        /product="putative membrane protein:(fragment)."
                        /protein_id="CAC03622.1"
                        /db_xref="GI:9843814"
                        /translation="IIGFGNTAFLVGSVCFLPSPSLERIALNLFVLGSLGWFVGSIGE
                        VFNHAKSRRTQS"
     misc_feature       1..126
                        /gene="SC8E7.01"
                        /note="nominal overlap with cosmid St10B8A between bases
                        9208..9333."
                        261..596
                        /gene="SC8E7.02"
                        /gene="SC8E7.02"
                        /note="SC8E7.02, doubtful CDS, len: 111 aa."
                        /codon_start=1
                        /transl_table=11
                        /product="hypothetical protein SC8E7.02."
                        /protein_id="CAC03623.1"
                        /db_xref="GI:9843815"
                        /translation="MPTDFGSGPLLPGASSVGEGGQEAAYGDQTDDEVGVLEPS
                        GIIVSALIIATIAPAATAITAGEKPPTTVHEERGEAGRGDTPRPVDPSPDPS
                        LRRADSR"
                        544..978
                        /note="shares 74.65% identity in 434 nt overlap with
                        repeated sequence 13658..14090"
                        complement(1088..1819)
                        /gene="SC8E7.03c"
                        complement(1088..1819)
                        /gene="SC8E7.03c"
                        /note="SC8E7.03c, possible DNA-binding protein, len: 243
                        aa. Weakly similar to several including: Streptomyces
                        ambofaciens protein of unknown function found within the
                        SrmR regulatory locus TR:Q00510(EMBL:X63451) SrmX (239
                        aa), fasta scores opt: 167 z-score: 217.0 E(): 0.00013
                        30.9% identity in 220 aa overlap and Deinococcus
                        radiodurans TR:Q9RTS7(EMBL:AE002009)
                        N-methyl-transferase-related protein (249 aa), fasta
                        scores opt: 160 z-score: 208.0 E(): 0.00041 27.1% identity

```

```

in 214 aa overlap. Contains a putative helix-turn-helix
motif situated between residues 60..81 (+3.06 SD)."
/codon_start=1
/transl_table=11
/product="putative DNA-binding protein."
/protein_id="CAC03624.1"
/db_xref="GI:9843816"
/translation="MPEGKLTGASVAAFGDLVRLPPKARVLDCACGTGOLAVGLAA
LGLDVATDASAMARHTQLAEQHGYSRLTRVSWDELSDRLDGSPPFVFCVGN
LAHGAERAFDALTAMSLRGLHVLSTRTWELVRAGSLDVGPRLVRRGRDA
VVYNNQIEQRDEEHLLEIAVAQVTADGSLVLTKSERLSCWFFRQEELVSOLHVGLE
VESNTFPDAENYTVITAKERAQS"
complement(1827..2105)
/gene="SC8E7.04c"
complement(1827..2105)
/gene="SC8E7.04c"
/note="SC8E7.04c, unknown, len: 92 aa. Contains 2xTTA
/leucine codons at the C-terminus."
/codon_start=1
/transl_table=11
/product="hypothetical protein SC8E7.04c."
/protein_id="CAC03625.1"
/db_xref="GI:9843817"
/translation="MGADRGVDMEAFVDAALGVGNGLNGLEAGQRDRGGIESRRNC
PSGTEFOTCGVSKGSDSLGRARAPCTLPSSVGVGCGRLRDLPLGL"
complement(1830..1832)
/gene="SC8E7.04c"
/note="TTA /leucine codon, possible target for bldA
regulation."
/label=bldA
complement(1851..1853)
/gene="SC8E7.04c"
/note="TTA /leucine codon, possible target for bldA
regulation."
/label=bldA
complement(2108..2401)
/gene="SC8E7.05c"
complement(2108..2401)
/gene="SC8E7.05c"
/note="SC8E7.05c, unknown, len: 97 aa. Highly similar to
the N-terminal of Streptomyces coelicolor mini-circle
putative transposase for IS11 SW:YM3_STRCO(EMBL:X15942)
SC3C8.19 (414 aa), fasta scores opt: 239 z-score: 303.4
E(): 2e-09 44.6% identity in 92 aa overlap."
/codon_start=1
/transl_table=11
/product="hypothetical protein SC8E7.05c."
/protein_id="CAC03626.1"
/db_xref="GI:9843818"
/translation="MIYCGIAWAERSHDVALVDNDGOLLAKRHYTDDAAGYKIFLGLL
VEYDSEENPACRTASSENGRRCCTCAQAGSLRSPFRGGRGAGGAVGAPAA"
complement(2408..2412)
complement(3106..3738)
/gene="SC8E7.06c"
complement(3106..3738)
/gene="SC8E7.06c"
/note="SC8E7.06, possible TetR-family transcriptional
regulator, len: 210 aa. Similar to several other members
of the TetR-family e.g. Streptomyces glaucescens
SW:TCMR_STRGA(EMBL:M80674) tetracenomycin C
transcriptional repressor, TCMR (226 aa), fasta scores
opt: 187 z-score: 235.6 E(): 1.2e-05 26.4% identity in
197 aa overlap and Streptomyces coelicolor
TR:CA888457(EMBL:AL353815) putative TetR-family
transcriptional regulator, SC86.28 (215 aa), fasta scores
opt: 239 z-score: 298.6 E(): 3.7e-09 32.0% identity in
178 aa overlap. Contains a Pfam match to entry PF00440
TetR, Bacterial regulatory proteins, tetR family with the
putative helix-turn-helix motif situated between residues
36..57 (+5.84 SD)."
/codon_start=1
/transl_table=11
/product="putative TetR-family transcriptional regulator."

```

gene	/label=SCJ30.03c		
	/product="hypothetical protein"		
	/protein_id="CAB53298.1"		
	/db_xref="GI:5763918"		
	/db_xref="SPTREMBL:O9SIY6"		
	/translation="MACVGGEDEWVYRVACQAGSQSAVDEGEGLAGVAVWVPRGG ECAGEGGFRVDHVGHQGGGAARGQAGVGKRAEPGAGVPAVGEMLLVVYPAGGD SARGPGGEGAWP"		
	complement(1760..2095)		
	/gene="SCJ30.04c"		
	complement(1760..2095)		
	/gene="SCJ30.04c"		
CDS	/note="SCJ30.04c, improbable CDS, function unknown, len: 111aa; predicted by GC frameplot and amino acid usage."		
	/codon_start=1		
	/transl_table=11		
	/label=SCJ30.04c		
	/product="hypothetical protein"		
	/protein_id="CAB53299.1"		
	/db_xref="GI:5763919"		
	/db_xref="SPTREMBL:O9SIY5"		
	/translation="MCGRRPEFTQLSCCCOLLGELVPPRGVRYVPTOFHDYRLHPA VDHLEFAAARNRMAFPEDPLCESRSTLSLLRLSRALRATSCRTGQRCFGGLDT QCQGPLLL"		
	2221..2265		
repeat_region	/note="Inverted and repeated at 3791-3835bp."		
	2509..3600		
	/gene="SCJ30.05"		
	2509..3600		
	/gene="SCJ30.05"		
	/note="SCJ30.05, unknown, len: 362aa; region of similarity to TR: CAB41280 (EMBL:AL049707) putative transferase from Streptomyces coelicolor (448 aa) fasta scores; Opt: 411, z-score: 409.4, E(): 1.9e-15, (32.8% identity in 274 aa overlap). Contains leucine tta codon, possible target for bldA regulation. Note large overlap with downstream CDS:"		
	/codon_start=1		
	/transl_table=11		
	/label=SCJ30.05		
	/product="hypothetical protein"		
gene	/protein_id="CAB53300.1"		
	/db_xref="GI:5763920"		
	/db_xref="SPTREMBL:O9SIY4"		
	/translation="MTGRTLFTGLATATCGSLLLKHFVAPGVPVGHIPWVLRLLI MLGTEGGVLFVAFRRVVRVRRHVRVLTGFELRGERYILYLRPFALDIRMSLPP PEAPGWMRSPLPELGTIMDFLVRQFTRHGVVAVGEGEELPLIGAORGYLPLEGM VERVELIOGASHVLSVAPGCTWEFFEALTMPERLYLVWCCGPEYDAFRSA VEKYAVRKSEBGSWTAPLPDPCFARPAKREWSPLRAFVTFDQWQPSLHWFF VTVPRIHVTMRRLVRIDALIGAWAALPQQAQSGTIPPPAPVYVTTTTPPPPVPLSP LPQQLGSTVGLNVRRPPTRRRGRQ"		
	complement(3357..4166)		
	/gene="SCJ30.06c"		
	complement(3357..4166)		
	/gene="SCJ30.06c"		
	/note="SCJ30.06c, unknown, len: 269aa; predicted amino acid usage. Note large overlap with downstream CDS."		
	/codon_start=1		
CDS	/transl_table=11		
	/label=SCJ30.06c		
	/product="hypothetical protein"		
	/protein_id="CAB53301.1"		
	/db_xref="GI:5763921"		
	/db_xref="SPTREMBL:O9SIY3"		
	/translation="MHRLETPRAAREGGEASIPITLHRAIRDLTPGERTGLAGGER EACKHDEFLARTGWEPGVGDPRHADPGACRCDGKPPRPSITWTFDCATNATGVAV TPGHPSRKSWLAALSRDUGDRPHRCGRARFVNRWQLNAAKFEAGAFQRIPLAY AHSEKAWAALGEGRDLPPTHVEQWVPVNFHCRPRRRLVRSGLTSPPTVLPSRCGCC SGRTGGGGVWTGAGGGIVPGEACRCGRAAHAPTNASMRSTSLRMVYT"		
	3403..3405		
	/gene="SCJ30.05"		
	/note="Leucine tta codon, possible target for bldA regulation."		
misc_feature	3791..3835		
	/repeat_unit		

/note="Inverted and repeated at 2221-2265bp. "
complement(4187..4597)
/gene="SCJ30.07c"
complement(4187..4597)
/gene="SCJ30.07c"
/note="SCJ30.07c, unknown, len: 136aa; predicted by GC
Frameplot, Hidden Markov model and amino acid usage."
/codon_start=1
/transl_table=11
/label="SCJ30.07c"
/product="hypothetical protein"
/protein_id="CA853302.1"
/db_xref="GI:5763922"
/db_xref="SPTREMBL:Q9SIY2"
/translation="MEACAPWRVTSISAGEARERDAAGLPHVVVYRFGAPLEVRVLS
VDLHVHGLWAYDAQDRRSHDLMDRLDDPSRLRLSVSWHCTGPETAEDKAGPRIT
VDMHQRLHGLGVLTGLRAPPPSLRPLPHR"
complement(4882..5172)

Query Match 74.2% Score 17.8; DB 1; Length 11311;
Best Local Similarity 90.5%; Pred. No. 7.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ggtgcagctagctcgaggggg 23
|||||iiiiiiiiiiiiiii
Db 2151 GGTGGCGTACGTCGAGGGGG 2131

RESULT 10

SC8E7
LOCUS Streptomyces coelicolor cosmid 8E7. linear BCT 16-AUG-2000
DEFINITION Streptomyces coelicolor cosmid 8E7.
ACCESSION AL391338
VERSION AL391338.1 GI:9843813
KEYWORDS ATP/GTP-binding protein; bldA; DNA-binding; insertion element
transposase; membrane; Na⁺/H⁺ antiporter; oxidoreductase;
prolipoprotein disacylglycerol transferase; pseudogene; repeat;
TetR-family transcriptional regulator.
SOURCE Streptomyces coelicolor A3(2).
ORGANISM Streptomyces coelicolor A3(2).
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE 1 (bases 1 to 39741)
AUTHORS Redenbach,M., Kieser,H.M., Denepaite,D., Eichner,A., Cullum,J.,
Kinashi,H. and Hopwood,D.A.
TITLE A set of ordered cosmids and a detailed genetic and physical map
for the 8 Mb Streptomyces coelicolor A3(2) chromosome
Mol. Microbiol. 21 (1), 77-96 (1996)
97000351
REFERENCE 2 (bases 1 to 39741)
AUTHORS Brown,S.P. and Harris,D.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 39741)
AUTHORS Thomson,N.R., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2000) Streptomyces coelicolor sequencing project,
Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
Colney, Norwich, Norfolk NR4 7UH, UK
Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded
by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are
available on the World Wide Web.
(URL: http://www.sanger.ac.uk/Projects/S_coelicolor/)
CDS are numbered using the following system eg SC7B7.01c. SC (S.
coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for

JOURNAL

Patent: JP 1999225770-A 334 24-AUG-1999;

COMMENT

NOVARTIS AG
OS PAG1265UP
PN JP 1999225770-A/334
PD 24-AUG-1999
PF 05-JAN-1998 JP 1998076818
PR 31-DEC-1996 CH 16/97
PI PETER PHILLIPSEN, REINER POHLMANN, SABINE STEINER, CHRISTINE MORE, PI JUERGEN WENDLAND, PHILIP KUNEHITTOU, CORINNE REBISHUN PC
(C12N15/09, C12N1/15, C12O1/18, C12Q1/68)/(C12N15/09, C12R1:645), PC
(C12N15/09, C12R1:865), (C12N1/15, C12R1:645), C12N15/00, PC
(C12N15/00, C12R1:645),
PC (C12N15/00, C12R1:865)
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..820
FT Location/Qualifiers
FT 1..820 /organism="PAG1265UP",
/organism="unidentified"
/db_xref="taxon:32644"

FEATURES

source

BASE COUNT 136 a 265 c 220 g 163 t 36 others

ORIGIN

Query Match 74.2%; Score 17.8; DB 6; Length 820;

Best Local Similarity 90.5%; Pred. No. 1.1e+03;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 gtcgacgtacgtcgagggggg 24

| | | | | | | | | | | | | | | | | | | | | |

Db 196 GTCGACGTACGTGCGGGGG 216

RESULT

9

SCJ30/c

LOCUS

SCJ30 Streptomyces coelicolor cosmid J30. 11311 bp DNA linear BCT 21-JUN-2000

DEFINITION

AL109973.1 GI:5763915

b1dA; Na+/H+ antiporter; transposase.

Streptomyces coelicolor A3(2).

Streptomyces coelicolor A3(2).

KEYWORDS

SOURCE

ORGANISM

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

1 (bases 1 to 11311)

Redenbach, M., Kieser, H. M., Denapaita, D., Eichner, A., Cullum, J.,

Kinashi, H. and Hopwood, D. A.

A set of ordered cosmids and a detailed genetic and physical map

for the 8 Mb Streptomyces coelicolor A3(2) chromosome

Mol. Microbiol. 21 (1), 77-96 (1996)

97000351

2 (bases 1 to 11311)

Sanders, D. C. and Harris, D.

Unpublished

3 (bases 1 to 11311)

Bentley, S. D., Parkhill, J., Barrell, B. G. and Rajandream, M. A.

Direct Submission

Submitted (20-AUG-1999) Streptomyces coelicolor sequencing project,

Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge.

CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.

David A. Hopwood, [3] John Innes Centre, Norwich Research Park,

Colney, Norwich, Norfolk NR4 7UH, UK

Notes:

Streptomyces coelicolor sequencing at The Sanger Centre is funded

by the BBSRC and Beowulf Genomics

Details of S. coelicolor sequencing at the Sanger Centre are

available on the World Wide Web

(URL: http://www.sanger.ac.uk/Projects/S_coelicolor/)

CDS are numbered using the following system eg SCJ30.01c, SC (S.

coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary

strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous.

The length in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for

CDS which show significant similarity to other CDS in the database.

The position of possible ribosome binding site sequences are given

where these have been used to deduce the initiation codon.

Gene prediction is based on positional base preference in codons

using a specially developed Hidden Markov Model (Krogh et al.,

Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot

program of Bibb et al., Gene 30:157-66(1984) as implemented at

<http://www.nih.gov/jp/jun/cgi-bin/frameplot.pl>

CAUTION: We may not have predicted the

correct initiation codon. Where possible we choose an initiation

codon (atg, gtg, ttg or att) which is preceded by an upstream

ribosome binding site sequence (optimally 5-13bp before the

initiation codon). If this cannot be identified we choose the most

upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the

sequenced clone. It may be shorter because we only sequence

overlapping sections once, or longer, because we arrange for a

small overlap between neighbouring submissions.

Cosmid J30 lies at the chromosome end and overlaps with cosmid J4

on the AseI-J genomic restriction fragment.

FEATURES

source

1..11311

/organism="Streptomyces coelicolor A3(2)"

/strain="A3(2)"

/db_xref="taxon:100226"

/clone="cosmid J30"

1..553

/gene="SCJ30.01"

<1..553

/gene="SCJ30.01"

/note="SCJ30.01, partial CDS, unknown, len: >183aa; weakly

similar to the C-terminal half of suspected transposition

related nucleotide binding proteins eg. TR:Q57461

(EMBL:U49101) TnIBdelta from plasmid pVsl (286 aa) fasta

scores: opt: 133, z-score: 170.1, E(): 0.041, (29.7%

identity in 148 aa overlap)."

/codon_start=2

/transl_table=11

/label="SCJ30.01

/product="hypothetical protein"

/protein_id="CAB53296.1"

/db_xref="GI:5763916"

/db_xref="SPTREMBL:Q9SIY8"

/translation="ITGAVCHTVAARVOLVMIIDEIHLNPRITTTGAQSADLIKDLT

RIGATFVYAGIDVTTPFTFTGVRGAOLAGRASLIDCAAFPSRLGDQPPFRDLITAMES

ALDLRHHPGTLPLAPLPHYHRTAGRIGSLARLQRAITALLDGTERTITKTTLDIAIR

LDHLAEQHYQPRAPRSRNTSTP"

complement(697..1143)

/gene="SCJ30.02c"

complement(697..1143)

/gene="SCJ30.02c"

/note="SCJ30.02c, unknown, len: 148aa"

/codon_start=1

/transl_table=11

/label="SCJ30.02c

/product="hypothetical protein"

/protein_id="CAB53297.1"

/db_xref="GI:5763917"

/db_xref="SPTREMBL:Q9SIY7"

/translation="MCVVRGDKLHGPAGCGSEPRVGGCLLVANAHQPLWRRVGPQG

VGVVAAGAAGAGADLLFLEPLPVRGGLGAVGVARQGEVGGQGVGEWESGEG

GGAPGRVIVIEFEGCQWAGVADAVKELQRLVGVGEPLP"

complement(1153..1503)

/gene="SCJ30.03c"

complement(1153..1503)

/gene="SCJ30.03c"

/note="SCJ30.03c, unknown, len: 116aa"

/codon_start=1

/transl_table=11

gene

CDS

gene

CDS

gene

CDS


```
CDS
complement(4781..5959)
/genes="dip"
/notes="Dip"
/codon_table=1
/product="pantothenate metabolism flavoprotein"
/protein_id="AG19087.1"
/db_xref="GI:10580167"
/translation="MLSGYNVAVGVGTSGAAVKKVVFVHLLRRRGACVRVMTESAQG
IHPWAVEFATNPVVTETGVPHVLCGRGMDVFFVAPATANTVGKIAAIVDDSD
PVYTCVTAVGADVPVVPVVPVVPVVPVVPVVPVVPVVPVVPVVPVVPVVPVVPV
PRESTIVHTAAGAEOPVVPVVPVVPVVPVVPVVPVVPVVPVVPVVPVVPVVPVVPV
VPGARTLVHDSAGDCAVYADYRDVSSAEMTATLTCACADALVSAALISDIYVE
ADKRLRGLDGLAVVANDAGVNDTRALFVTDVSVSEYAGHKGRLGARIABLAFA
"
complement(6567..7259)
/genes="VNG0573C"
/notes="VNG0573C"
/notes="conserved hypothetical protein"
/codon_start=1
/codon_table=1
/product="Vng0573c"
/protein_id="AG19088.1"
/db_xref="GI:10580168"
/translation="MGGSVPPMISVLRRTCGRLSVAASGRVANGQTKVFSSRDS
LAARNSETDVGSRVLRGGVAAAAGAAATGTAAGGEGGDDGGEKVVKVPNGS
NVFKPAEMTVKPGTKVRFVWESGHNHATEVPGDADGCVSTDIAQPPKEYHTFDGP
TGEYNVCTPHASLGMKGITIVTDSPPENKGYQTIVPDSAKTMVAGMSMTSVLGLA
VFEMKYGGDYG"
7258..7515
/genes="VNG0574C"
/notes="VNG0574C"
/notes="conserved hypothetical protein"
/codon_start=1
/codon_table=1
/product="Vng0574c"
/protein_id="AG19089.1"
/db_xref="GI:10580169"
/translation="NSTVESLREELTDAFADYVPVKQNMGLVPALPSPGSPSPFAG
GESWTAMELATKLSSRADFPYDSVDTLVDDVIAGLKDEDEL"
7566..8867
/genes="Ywad"
/notes="Ywad"
/notes="Ywad"
/codon_start=1
/codon_table=1
/product="aminopeptidase"
/protein_id="AG19090.1"
/db_xref="GI:10580170"
/translation="WTDWIGDFTSTGTHLQRLVDTPTRWAGSDGEREAARTRDA
LAQHADDALWLDTFVQGWTRGDSAIETPADTTADTIALPRSPAGTAAGFEVDLSGLP
GFEDRDLGALVVAADVPDWDYDRHLHREKRYAAVAAAGAFVYMNHVAGCLPATG
SVYTGDDPGEIPAVGVSTETGRLSRFEHDTVTLVDADTHDATSQNVHATLGPDT
DAELLVTSVDAHDITDGDNDGACGTAMAVELARILAARESALDTRVHFVCFGEVEG
75.8%; Score 18.2; DB 1; Length 10293;
Query Match
Best Local Similarity 87.0%; Pred. No. 5.2e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gggctgcagctacgcagggggg 24
|||||
Db 1784 GCCTGCGGTACGTCGAGGCGG 1762

RESULT 6
A85676
LOCUS
DEFINITION
Sequence 335 from Patent EP0866129.
AUTHORS
Juergeen, W.P.K.K. and Rebishun.
TITLE
Genome DNA sequence of Ashbya gossypii and use thereof.

A85676.1 GI:6734275
Eremothecium gossypii.
Eremothecium gossypii.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Eremotheciaceae; Eremothecium.
REFERENCE
1 (bases 1 to 820)
Mohr, C. and Knechtle, P.
Genomic DNA sequences of Ashbya gossypii and uses thereof
Patent: EP 0866129-A 335 23-SEP-1998;
CIBA GEIGY AG (CH)
FEATURES
Location/Qualifiers
source
1..820
/organism="Eremothecium gossypii"
/db_xref="taxon:33169"
BASE COUNT 136 a 265 c 220 g 163 t 36 others
ORIGIN
Query Match
Best Local Similarity 74.2%; Score 17.8; DB 6; Length 820;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 gtcgacgtacgtcagggggg 24
|||||
Db 196 GTCGACGTACGTTCCGGGGG 216

RESULT 7
A85676
LOCUS
DEFINITION
Sequence 335 from patent US 6239264.
AUTHORS
Philippsen, P., Pohlmann, R., Steiner-Lange, S., Mohr, C., Wendland, J.,
Knechtle, P. and Reibischung, C.
Genomic DNA sequences of Ashbya gossypii and uses thereof
Patent: US 6239264-A 335 29-MAY-2001;
CIBA GEIGY AG (CH)
FEATURES
Location/Qualifiers
source
1..820
/organism="unknown"
/db_xref="taxon:33169"
BASE COUNT 136 a 265 c 220 g 163 t 36 others
ORIGIN
Query Match
Best Local Similarity 74.2%; Score 17.8; DB 6; Length 820;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 gtcgacgtacgtcagggggg 24
|||||
Db 196 GTCGACGTACGTTCCGGGGG 216

RESULT 8
E65694
LOCUS
DEFINITION
Genome DNA sequence of Ashbya gossypii and use thereof.
AUTHORS
Juergeen, W.P.K.K. and Rebishun.
TITLE
Genome DNA sequence of Ashbya gossypii and use thereof.

E65694
LOCUS
DEFINITION
Genome DNA sequence of Ashbya gossypii and use thereof.
AUTHORS
Juergeen, W.P.K.K. and Rebishun.
TITLE
Genome DNA sequence of Ashbya gossypii and use thereof.
```

Db 706 GGGTCCAGCGTGGTCAGGGG 727

RESULT 5

AE005007/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AE005007 10293 bp DNA linear BCT 12-FEB-2001
Halobacterium sp. NRC-1 section 38 of 170 of the complete genome.
AE005007 AE004437
AE005007.1 GI:10580159

Halobacterium sp. NRC-1.

Halobacterium sp. NRC-1

Archaea; Euryarchaeota; Halobacteria; Halobacteriales;

Halobacteriaceae; Halobacterium.

1 (bases 1 to 10293)

Ng,W.V., Kennedy,S.P., Mahairas,G.G., Berquist,B., Pan,M.,

Shukla,H.D., Lasky,S.R., Balliga,N., Thorsson,V., Sbrogna,J.,

Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Welter,R., Goo,Y.A.,

Leithausen,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W.,

Maddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H.,

Isenbarger,T.A., Peck,R.F., Pohlshrod,M., Spudich,J.L.,

Jung,K.-H., Alam,M., Freitas,T., Hou,S., Daniels,C.J., Dennis,P.P.,

Omer,A.D., Ehardt,H., Lowe,T.M., Liang,P., Riley,M., Hood,L. and

DasSarma,S.

From the cover: genome sequence of halobacterium species NRC-1

Proc. Natl. Acad. Sci. USA 97 (22), 12176-12181 (2000)

11016950

2 (bases 1 to 10293)

Ng,W.V., Kennedy,S.P., Mahairas,G.G., Berquist,B., Pan,M.,

Shukla,H.D., Lasky,S.R., Balliga,N., Thorsson,V., Sbrogna,J.,

Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Welter,R., Goo,Y.A.,

Leithausen,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W.,

Maddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H.,

Isenbarger,T.A., Peck,R.F., Pohlshrod,M., Spudich,J.L.,

Jung,K.-H., Alam,M., Freitas,T., Hou,S., Daniels,C.J., Dennis,P.P.,

Omer,A.D., Ehardt,H., Lowe,T.M., Liang,P., Riley,M., Hood,L. and

DasSarma,S.

Direct Submission

Submitted (14-JUL-2000) Institute for Systems Biology, 4225

Roosevelt Way NE, Seattle, WA 98105, USA

Location/Qualifiers

1..10293

/organism="Halobacterium sp. NRC-1"

/strain="NRC-1"

/db_xref="taxon:64091"

complement(104..1246)

/gene="ndhc2"

/note="ndhc2"

complement(104..1246)

/gene="ndhc2"

/note="ndhc2"

/codon_start=1

/transl_table=11

/product="NADH dehydrogenase/oxidoreductase"

/protein_id="AAG19080.1"

/db_xref="GI:10580160"

/translation="MLVAGLSGMTVTADLFNWMYVFIETGLTAYALVAAGDSRSIV

AALKYLVGTGSLFLIGYAVGTGLNMDLAADLADADPVLQAQAFALIVGL

FIKVAFLPQPDYAGADPTVSALISALYSTVAYALLRVYTVFGAGFLDSTPV

ARAILIGAVSVIGSLLAVSQSEIKRVLAYSQSGFLILAAIAGNDTALGAAL

HLVGHVMKGLFLTAGVATETGARTIDFGLADSPVAAGAFGLAYSMVGVPPT

VGFAGKWYIAGAAEAGSWALLAVIVASTLTLLAYFLRLVFERFRAPIVADGRSD

GAETPASSGTGATVVGNTVCAVALGFGAFVSTQKAILTQLLS"

complement(1578..1934)

/gene="VNG0565C"

complement(1578..1934)

/gene="VNG0565C"

/note="conserved hypothetical protein"

/codon_start=1

/transl_table=11

/product="Vng0565c"

/protein_id="AAG19081.1"

/db_xref="GI:10580161"

gene

CDS

/translation="MMEALAAKDYVIAFLLIGTGYVYVMIAGNVLKVKVIGMNFQIG
IFLEPIASAYVEGGTPPLTSETPHVSPVPHVLLITAIWGVALTAVAGLMIVRIYDD
YGTIDERTLKEVRSSD" .2398)
/gene="VNG0564H"
complement(1931..2398)
/gene="VNG0564H"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng0564h"
/protein_id="AAG19082.1"
/db_xref="GI:10580162"
/translation="MSDRPSSQROSYIESIIMGTVRVVSFPVFTYGLFVMEHGA
AGGFGGQVIVATVMTMLGIAFGIDPVREWGERLLVIAVGGVTAFAVASSTVLLG
GAFIDVSFVGGHASKYSIEIIEFIGAIVASTITGLFIAIDAGTSDDGNDK"
complement(2395..2925)
/gene="VNG0566C"
complement(2395..2925)
/note="conserved hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng0566c"
/protein_id="AAG19083.1"
/db_xref="GI:10580163"
/translation="MNIEPLLVFVATATATTVLRDLVASIIAFATYSLGVAVVVV
LQAPDVGLTEAAVAGAGVTVLFTIAKTVRPGDEAFERLNVPAFGAAVLVGLLT
VTHALPTIGDPSNIAVTSEVTRYIDHAYHEAGVKNVTVLASYRGDTLGEAVVY
SAGVGLIVLREVEFA"
complement(2922..3029)
/gene="VNG0570H"
complement(2922..3029)
/note="VNG0570H"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng0570h"
/protein_id="AAG19084.1"
/db_xref="GI:10580164"
/translation="MFINPTAAHIAARAADQGDIDPWTADDHEGGDGE"
complement(3245..3427)
/gene="VNG0568C"
complement(3245..3427)
/note="VNG0568C"
/note="conserved hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng0568c"
/protein_id="AAG19085.1"
/db_xref="GI:10580165"
/translation="MQDRVIGVNAIGTNVIVILVGAATETALDIAIVALLNFL
MSIAISKFTVERGGVL"
complement(3517..4611)
/gene="VNG0571C"
complement(3517..4611)
/note="VNG0571C"
/note="conserved hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng0571c"
/protein_id="AAG19086.1"
/db_xref="GI:10580166"
/translation="MADNDLLIPVADSETVQTVSHAADRAQERDGTPTLHFVCPVSE
RPNADASAAAEEALDRVSVMTDGLDGDASRAVETSLVGLDVLFPNGDYADY
FGAYARENGVTAVLDPRNPLGNAPLLPLPERLEAAGLTTELAPVERATRRGPITR
SAGIGQFALLFGAAFGVOLITELTTLNLTITGATITATVAGVLRVSTQGSVRPRL
VGRVGRMLVAPVLLLEIAKANVAIAKVLHPLRPIDPAVVEFDAVMSLPATTLAN
SITLPGTLIVDVTQRHTVHSLTADARLLAGGLERAVRFVYGRAMRSATPSE
GETPDGSVDSAGIDSASASRTDNTNAGDSE"
complement(4781..5959)
/gene="dfp"
/note="VNG0572C"

gene

Query Match 100.0%; Score 24; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcagctacgtcagggggg 24
|||||
Db 1 GGGGTCGACGTACGTGAGGGGG 24

RESULT 2
AX104852
LOCUS AX104852 24 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1044 from Patent WO0122972.
ACCESSION AX104852
VERSION AX104852.1 GI:13921049
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 24)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 1044 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)

FEATURES
source Location/Qualifiers
1..24
/organism="synthetic construct"
/db_xref="taxon:32630" 3 t

BASE COUNT 3 a 4 c 14 g 3 t
ORIGIN

Query Match 100.0%; Score 24; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcagctacgtcagggggg 24
|||||
Db 1 GGGGTCGACGTACGTGAGGGGG 24

RESULT 3
AX105127
LOCUS AX105127 24 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 25 from Patent WO0122990.
ACCESSION AX105127
VERSION AX105127.1 GI:13921277
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 24)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 25 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)

FEATURES
source Location/Qualifiers
1..24
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
misc_feature 1..2
/note="Backbone has phosphorothioate linkages."
misc_feature 3..18
/note="Backbone has phosphodiester linkages."
misc_feature 19..23
/note="Backbone has phosphorothioate linkages."
misc_feature 24
/note="Backbone has phosphodiester linkages."

BASE COUNT 3 a 4 c 14 g 3 t
ORIGIN

Query Match 100.0%; Score 24; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcagctacgtcagggggg 24
|||||
Db 1 GGGGTCGACGTACGTGAGGGGG 24

RESULT 4
BTFT2GEN
LOCUS BTFT2GEN 1738 bp DNA linear MAM 02-OCT-1997
DEFINITION Bos taurus FUT2 gene.
ACCESSION X99620
VERSION X99620.1 GI:2464960
KEYWORDS FUT2 gene.
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 1738)
AUTHORS Petit,J.M.
TITLE Bovine type Se FUT2 gene: partial sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1738)
AUTHORS Petit,J.M.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-1996) J.M. Petit, Universite de Limoges, Institut
de Biotechnologie, 123 Avenue Albert Thomas, 87060 Cedex, Limoges,
FRANCE
REMARK Revised by [3]
REFERENCE 3 (bases 1 to 1738)
AUTHORS Petit,J.M.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-1997) J.M. Petit, Universite de Limoges, Institut
de Biotechnologie, 123 Avenue Albert Thomas, 87060 Cedex, Limoges,
FRANCE
COMMENT On Oct 4, 1997 this sequence version replaced gi:1480079.

FEATURES
source Location/Qualifiers
1..1738
/organism="Bos taurus"
/db_xref="taxon:9913"
/dev_stage="adult"
115..1149
/gene="FUT2"
115..1149
/gene="FUT2"
/codon_start=1
/protein_id="CAA67931.1"
/db_xref="GI:2464961"
/translation="MFSTQTFPTAPFLFVFTASTIFHLHQLRLEKMQPTWELAL
EPATMTPSRPQRPQLKGMWTINATIRLGNQMGVEATLYALAKMNGRAAFIPQMHS
TLAIFRIFLVLHDATARSVPWQVHNDWMEQYRHPGVEYRLTGYPCSWTFYHH
LRAEILQETLHAHVREAAQNLGLRVNSRPSTVGVVRRGDYVHVMPNVKGVV
ADRRYLEQALDHFRRARYSAPFVSSNGHNCWRENINASRGDVFVAGNNGSPAKDF
ALLTCNHTMTITGTFGIWAAYLAGGETIYLTLPDPSFLKIFRPAFLPKWIGI
PADLSPLKKH"

BASE COUNT 389 a 534 c 438 g 376 t 1 others
ORIGIN

Query Match 78.3%; Score 18.8; DB 4; Length 1738;
Best Local Similarity 90.9%; Pred. No. 3.6e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggtcagctacgtcaggggg 22
|||||

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:58:20 ; Search time 2778.35 seconds
(without alignments)
180.768 Million cell updates/sec

Title: US-09-672-126-25

Perfect score: 24

Sequence: 1 ggggtcgactgactcgaggggg 24

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.on.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg_hum.*

31: em.htg_inv.*

32: em.htg_Other.*

33: em.htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

	1	24	100.0	24	6	AX104782
AX104782	1	24	100.0	24	6	AX104782
LOCUS	2	24	100.0	24	6	AX104852
DEFINITION	3	24	100.0	24	6	AX105127
ACCESSION	4	18.8	78.3	1738	4	BFU2G2GEN
VERSION	5	18.2	75.8	10293	1	AE005007
KEYWORDS	6	17.8	74.2	820	6	AB5676
SOURCE	7	17.8	74.2	820	6	AR155169
ORGANISM	8	17.8	74.2	820	6	E65694
REFERENCE	9	17.8	74.2	11311	1	SCJ30
AUTHORS	10	17.8	74.2	39741	1	SC8E7
TITLE	11	17.8	74.2	158580	2	AC095115
JOURNAL	12	17.6	73.3	36028	1	SCE41
FEATURES	13	17.6	73.3	36539	1	SGK7
source	14	17.6	73.3	65172	2	AC104910
	15	17.6	73.3	77506	2	AC096238
	16	17.6	73.3	100773	8	AF468201
	17	17.6	73.3	104134	2	AC103443
	18	17.6	73.3	110000	2	LMFLCHR31_12
	19	17.6	73.3	112855	2	AP003608
	20	17.6	73.3	135957	2	AC099271
	21	17.6	73.3	171069	2	AC103141
	22	17.6	73.3	189290	2	AC106165
	23	17.6	73.3	241714	2	AC079564
	24	17.4	72.5	249262	2	AC079430
	25	17.2	71.7	821	9	AF136645
	26	17.2	71.7	2068	3	PAC130877
	27	17.2	71.7	2487	1	SEU58990
	28	17.2	71.7	32704	1	SC8A11
	29	17.2	71.7	38543	1	SCBAC16H6
	30	17.2	71.7	116578	2	AC094735
	31	17.2	71.7	152294	2	AC094504
	32	16.8	70.0	11034	1	AE005065
	33	16.8	70.0	16766	6	AX347059
	34	16.8	70.0	134213	9	CNS05TEV
	35	16.6	69.2	339	3	AF322489
	36	16.6	69.2	899	9	HSERC55D
	37	16.6	69.2	1233	14	AF298586
	38	16.6	69.2	1357	8	AF302806
	39	16.6	69.2	1845	14	SH1EPO
	40	16.6	69.2	2282	10	RATRGER4
	41	16.6	69.2	2440	8	AF421889
	42	16.6	69.2	5487	6	AX346501
	43	16.6	69.2	6242	6	AX347050
	44	16.6	69.2	8250	8	AF380365
	45	16.6	69.2	8438	14	SH1LLT

ALIGNMENTS

RESULT	1	AX104782	Sequence 974 from Patent WO0122972	24 bp	DNA	linear	PAT 30-APR-2001
LOCUS	AX104782	Sequence 974 from Patent WO0122972					
DEFINITION	AX104782	Sequence 974 from Patent WO0122972					
ACCESSION	AX104782	Sequence 974 from Patent WO0122972					
VERSION	AX104782.1	GI:13920979					
KEYWORDS		synthetic construct					
SOURCE		synthetic construct					
ORGANISM		artificial sequence					
REFERENCE		1 (bases 1 to 24)					
AUTHORS		Krieg, A.M., Schetter, C. and Vollmer, J.C.					
TITLE		Immunostimulatory nucleic acids					
JOURNAL		Patent: WO 0122972-A 974 05-APR-2001					
FEATURES		UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical GmbH (DE)					
source		Location/Qualifiers					
		1. .24					
		/organism="synthetic construct"					
		/db_xref="taxon:32630"					
BASE COUNT		3 a 4 c 14 g 3 t					
ORIGIN							

THIS PAGE BLANK (USPTO)

Job time: 16039 sec

1

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1320 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1320
US-09-031-606-8

Query Match 76.0%; Score 15.2; DB 3; Length 1320;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
||| ||||| ||||| |||
Db 897 GGTGGCACCGGTGAGGGTGG 878

RESULT 14

US-08-461-775-10/c
Sequence 10, Application US/08461775
Patent No. 5858773
GENERAL INFORMATION:
APPLICANT: MAZODIER, Philippe
APPLICANT: GUGLIEMI, Gerard
TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,775
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/050,313
FILING DATE: 10-MAY-1993
APPLICATION NUMBER: FR 9011186
FILING DATE: 10-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 010830-035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1620
US-08-461-775-10

Query Match 76.0%; Score 15.2; DB 2; Length 1620;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
||| ||||| ||||| |||
Db 897 GGTGGCACCGGTGAGGGTGG 878

RESULT 15

US-09-031-606-10/c
Sequence 10, Application US/09031606
Patent No. 6153404
GENERAL INFORMATION:
APPLICANT: MAZODIER, Philippe
APPLICANT: GUGLIEMI, Gerard
TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,606
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/050,313
FILING DATE: 10-MAY-1993
APPLICATION NUMBER: FR 9011186
FILING DATE: 10-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 010830-035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1620
US-09-031-606-10

Query Match 76.0%; Score 15.2; DB 3; Length 1620;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
||| ||||| ||||| |||
Db 897 GGTGGCACCGGTGAGGGTGG 878

Search completed: August 10, 2002, 03:06:13

;; TITLE OF INVENTION: PANCREATIC DISEASE
;; NUMBER OF SEQUENCES: 24
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Banner & Witcoff, Inc.
;; STREET: One Financial Center
;; CITY: Boston
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02111
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WordPerfect 6.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/881,450A
;; FILING DATE: June 24, 1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kathleen M. Williams
;; REGISTRATION NUMBER: 34,380
;; REFERENCE/DOCKET NUMBER: 11275/7823
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-345-9100
;; TELEFAX: 617-345-9111
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 400 nucleotides
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: genomic DNA
;; FEATURE:
;; NAME/KEY: human IPF-1 gene
;; LOCATION: exon 1
US-08-881-450A-1

Query Match 76.0%; Score 15.2; DB 4; Length 400;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggggtcacccgtgagggggg 20
||||| ||| |||||

Db 250 GGGGTCGTCGCGAGGGGGG 231

RESULT 12
US-08-461-775-8/c
;; Sequence 8, Application US/08461775
;; Patent No. 5858773
;; GENERAL INFORMATION:
;; APPLICANT: MAZODIER, Philippe
;; APPLICANT: GUGLIEMI, Gerard
;; TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
;; TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
;; NUMBER OF SEQUENCES: 15
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burns, Doane, Swecker & Mathis
;; STREET: George Mason Bldg., Washington & Prince Sts.
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/461,775
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/050,313
;; FILING DATE: 10-MAY-1993
;; APPLICATION NUMBER: FR 9011186
;; FILING DATE: 10-SEP-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Crane-Feury, Sharon E
;; REGISTRATION NUMBER: 36,113
;; REFERENCE/DOCKET NUMBER: 010830-035
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2021
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1320 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..1320
US-08-461-775-8

Query Match 76.0%; Score 15.2; DB 2; Length 1320;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggggtcacccgtgagggggg 20
||||| ||| |||||

Db 897 GGTGGCAGGTGAGGGTGG 878

RESULT 13
US-09-031-606-8/c
;; Sequence 8, Application US/09031606
;; Patent No. 6153404
;; GENERAL INFORMATION:
;; APPLICANT: MAZODIER, Philippe
;; APPLICANT: GUGLIEMI, Gerard
;; TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
;; TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
;; NUMBER OF SEQUENCES: 15
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burns, Doane, Swecker & Mathis
;; STREET: George Mason Bldg., Washington & Prince Sts.
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/031,606
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/050,313
;; FILING DATE: 10-MAY-1993
;; APPLICATION NUMBER: FR 9011186
;; FILING DATE: 10-SEP-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Crane-Feury, Sharon E
;; REGISTRATION NUMBER: 36,113
;; REFERENCE/DOCKET NUMBER: 010830-035
;; TELECOMMUNICATION INFORMATION:

;; CITY: Rahway
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07065
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/04801
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Wallen III, John W.
;; REGISTRATION NUMBER: 35,403
;; REFERENCE/DOCKET NUMBER: 19179
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (908) 594-3905
;; TELEFAX: (908) 594-4720
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1700 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cdna
PCT-US95-04801-1

Query Match 79.0%; Score 15.8; DB 5; Length 1700;
Best Local Similarity 89.5%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gggtcacggtagggggg 20
|||||

Db 1048 GGGTCAACGTTGAGGTGG 1030

RESULT 9
US-08-386-063-1
; Sequence 1, Application US/08386063
; Patent No. 6008200
; GENERAL INFORMATION:
; APPLICANT: Arthur M. Krieg, M.D.
; TITLE OF INVENTION: IMMUNOMODULATORY OLIGONUCLEOTIDES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/386,063
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ARNOLD, BETH E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: UIZ-013CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs

;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA
US-08-386-063-1

Query Match 76.0%; Score 15.2; DB 3; Length 20;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggggtcacggtagggggg 20
|||||

Db 1 GGGTCAACGTTGAGGGGG 20

RESULT 10
US-08-386-063-1
; Sequence 1, Application US/08386063
; Patent No. 6194388
; GENERAL INFORMATION:
; APPLICANT: Arthur M. Krieg, M.D.
; TITLE OF INVENTION: IMMUNOMODULATORY OLIGONUCLEOTIDES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/386,063
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ARNOLD, BETH E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: UIZ-013CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-386-063-1

Query Match 76.0%; Score 15.2; DB 4; Length 20;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggggtcacggtagggggg 20
|||||

Db 1 GGGTCAACGTTGAGGGGG 20

RESULT 11
US-08-881-450A-1/c
; Sequence 1, Application US/08881450A
; Patent No. 6274310
; GENERAL INFORMATION:
; APPLICANT: Habener, J.F. and Stoffers, D.A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING

```

RESULT      8
PCT-US95-04801-1/c
; Sequence 1, Application PC/TUS9504801
; GENERAL INFORMATION:
; APPLICANT: Martih, Juan F.
; APPLICANT: Coque, Juan R.
; APPLICANT: Enuiga, Francisco J.
; APPLICANT: Fuente, Juan L.
; APPLICANT: Llaena, Francisco J.
; APPLICANT: Liras, Paloma
; TITLE OF INVENTION: DNA ENCODING
; TITLE OF INVENTION: LATE GENES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: P.O. Box 2000

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic oligonucleotide
US-09-030-701-63

Query Match 84.0%; Score 16.8; DB 4; Length 20;
Best Local Similarity 90.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
||||| || |||||
Db 1 ggggtcaacgttgagggggg 20

RESULT 3
US-08-960-774-90
; Sequence 90, Application US/08960774
; Patent No. 6239116
; GENERAL INFORMATION:
; APPLICANT: Krieg et al.
; TITLE OF INVENTION: IMMUNOSTIMULATORY NUCLEIC ACID MOLECULES
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,774
; FILING DATE: 30-October-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Serial No. 6239116 08/738,652
; FILING DATE: October 30, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 08918/012001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-960-774-90

Query Match 84.0%; Score 16.8; DB 4; Length 20;
Best Local Similarity 90.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
||||| || |||||
Db 1 GGGGTCAACGTTGAGGGGGG 20

RESULT 4
US-09-082-649B-52
; Sequence 52, Application US/09082649B
; Patent No. 6339068
; GENERAL INFORMATION:
; APPLICANT: Davis, Heather L.
; APPLICANT: Krieg, Arthur M.
; APPLICANT: Schorr, Joachim
; APPLICANT: Wu, Tong
; TITLE OF INVENTION: Vectors and Methods for Immunization or
; TITLE OF INVENTION: Therapeutic Protocols
; FILE REFERENCE: C1039/7009
; CURRENT APPLICATION NUMBER: US/09/082,649B
; CURRENT FILING DATE: 1998-05-20
; PRIOR APPLICATION NUMBER: US 60/047,233
; PRIOR FILING DATE: 1997-05-20
; PRIOR APPLICATION NUMBER: US 60/047,209
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic oligonucleotide
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Has a phosphorothioate backbone.
US-09-082-649B-52

Query Match 84.0%; Score 16.8; DB 4; Length 20;
Best Local Similarity 90.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
||||| || |||||
Db 1 ggggtcaacgttgagggggg 20

RESULT 5
US-09-082-649B-59
; Sequence 59, Application US/09082649B
; Patent No. 6339068
; GENERAL INFORMATION:
; APPLICANT: Davis, Heather L.
; APPLICANT: Krieg, Arthur M.
; APPLICANT: Schorr, Joachim
; APPLICANT: Wu, Tong
; TITLE OF INVENTION: Vectors and Methods for Immunization or
; TITLE OF INVENTION: Therapeutic Protocols
; FILE REFERENCE: C1039/7009
; CURRENT APPLICATION NUMBER: US/09/082,649B
; CURRENT FILING DATE: 1998-05-20
; PRIOR APPLICATION NUMBER: US 60/047,233
; PRIOR FILING DATE: 1997-05-20
; PRIOR APPLICATION NUMBER: US 60/047,209
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic oligonucleotide
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Has SOS-ODN backbone with two S-linkages at the 5'
; OTHER INFORMATION: end, five S-linkages at the 3' end, and O-linkages
; OTHER INFORMATION: in between.
US-09-082-649B-59

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:06:09 ; Search time 277.54 seconds
(without alignments)
17.701 Million cell updates/sec

Title: US-09-672-126-24

Perfect score: 20

Sequence: 1 ggggtcacccgtgagggggg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.8	84.0	20	4	US-08-738-652-12
2	16.8	84.0	20	4	US-09-030-701-63
3	16.8	84.0	20	4	US-08-960-774-90
4	16.8	84.0	20	4	US-09-082-649B-52
5	16.8	84.0	20	4	US-09-082-649B-59
6	16.4	82.0	3435	1	US-08-366-577-1
7	16.4	82.0	3435	5	PCT-US96-00005-1
8	15.8	79.0	1700	5	PCT-US95-04801-1
9	15.2	76.0	20	3	US-08-386-063-1
10	15.2	76.0	20	4	US-08-386-063-1
11	15.2	76.0	400	4	US-08-881-450A-1
12	15.2	76.0	1320	2	US-08-461-775-8
13	15.2	76.0	1320	3	US-09-031-606-8
14	15.2	76.0	1620	2	US-08-461-775-10
15	15.2	76.0	1620	3	US-09-031-606-10
16	15.2	76.0	1792	4	US-08-965-762-7
17	15.2	76.0	2167	2	US-08-461-775-9
18	15.2	76.0	2167	3	US-09-031-606-9
19	15.2	76.0	2668	2	US-08-461-775-11
20	15.2	76.0	2668	3	US-09-031-606-11
21	15.2	76.0	2886	1	US-08-073-383-3
22	15.2	76.0	2886	3	US-08-328-239A-2
23	15.2	76.0	2886	5	PCT-US94-06365-3
24	15.2	76.0	2886	5	PCT-US95-13661-2
25	15.2	76.0	2890	3	US-08-848-810-1
26	15.2	76.0	2940	1	US-08-428-415-3
27	15.2	76.0	2940	1	US-08-379-685-3

28 15.2 76.0 2940 2 US-08-854-029-3 Sequence 3, Appli
29 15.2 76.0 2940 4 US-08-428-762-3 Sequence 3, Appli
30 15.2 76.0 3070 1 US-08-428-732-3 Sequence 3, Appli
31 15.2 76.0 4926 2 US-08-853-310-1 Sequence 1, Appli
32 15.2 76.0 5658 4 US-08-881-450A-23 Sequence 23, Appli
33 15.2 76.0 6180 1 US-08-386-727-1 Sequence 1, Appli
34 15.2 76.0 6180 2 US-08-600-452A-1 Sequence 1, Appli
35 15.2 76.0 13987 2 US-08-804-227C-13 Sequence 13, Appli
36 15.2 76.0 43280 2 US-08-804-227C-1 Sequence 1, Appli
37 15.2 76.0 4403765 4 US-09-103-840A-2 Sequence 2, Appli
38 15.2 76.0 4411529 4 US-09-103-840A-1 Sequence 1, Appli
39 14.8 74.0 1218 3 US-09-012-072-1 Sequence 1, Appli
40 14.8 74.0 1218 4 US-09-120-601-1 Sequence 2, Appli
41 14.8 74.0 1889 2 US-09-026-587-2 Sequence 2, Appli
42 14.8 74.0 1889 2 US-09-227-420-2 Sequence 2, Appli
43 14.8 74.0 4200 1 US-07-841-654B-1 Sequence 1, Appli
44 14.8 74.0 4200 1 US-07-946-234A-1 Sequence 1, Appli
45 14.8 74.0 4200 1 US-08-123-161A-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-738-652-12
; Sequence 12, Application US/08738652B
; Patent No. 6207646
; GENERAL INFORMATION:
; APPLICANT: Krieg, Arthur M.
; TITLE OF INVENTION: Immunostimulatory Nucleic Acid Molecules
; FILE REFERENCE: C1039/7004 HCL
; CURRENT APPLICATION NUMBER: US/08/738,652B
; CURRENT FILING DATE: 1996-10-30
; EARLIER APPLICATION NUMBER: US 08/276,358
; EARLIER FILING DATE: 1994-07-15
; EARLIER APPLICATION NUMBER: US 08/386,063
; EARLIER FILING DATE: 1995-02-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
US-08-738-652-12

Query Match 84.0%; Score 16.8; DB 4; Length 20;
Best Local Similarity 90.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ggggtcacccgtgagggggg 20
||||| |||||
Db 1 ggggtcaccttgagggggg 20

RESULT 2
US-09-030-701-63
; Sequence 63, Application US/09030701B
; Patent No. 6214806
; GENERAL INFORMATION:
; APPLICANT: Krieg, Arthur M.
; APPLICANT: Schwartz, David A.
; TITLE OF INVENTION: USE OF NUCLEIC ACIDS CONTAINING UNMETHYLATED CpG DINUCLEOTIDE IN THE TREATMENT OF LPS-ASSOCIATED DISORDERS
; TITLE OF INVENTION: UNMETHYLATED CpG DINUCLEOTIDE IN THE TREATMENT OF LPS-ASSOCIATED DISORDERS
; FILE REFERENCE: C1039/7011
; CURRENT APPLICATION NUMBER: US/09/030,701B
; CURRENT FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/039,405
; PRIOR FILING DATE: 1997-02-28
; NUMBER OF SEQ ID NOS: 65

	Matches	18;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
Qy	1	ggggtcacccggtgagggggg	20							
Db	63	GGGTCACAGGGGAGGGGGG	44							

RESULT	15
BB593432	
LOCUS	
DEFINITION	BB593432 RIKEN full-length enriched, 4 days neonate male adipose Mus musculus cDNA clone B430002H17 5' , mRNA sequence.
	231 bp mRNA linear EST 30-NOV-2000

EST.	
KEYWORDS	Mus mouse.
SOURCE	Mus musculus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 231)
REFERENCE	Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P. Hanaoka,T., Hayatsu,N., Hiroaka,T., Hirozane,T., Hodoama,Y., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J., Kojima,Y., Konno H., Kusakabe,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Okazaki,Y., Okido,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Sato,K., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka T., Toyota,T., Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshiki,A., Muramatsu,M. and Hayashizaki.Y.
TITLE	RIKEN Mouse ESTs (Aizawa, K. et al. 2000) Riken Mammalian Genome Project

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermotabilization and thermotactivation of thermostable enzymes by trehalose and its application for the synthesis of full-length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

	Location/Qualifiers
1.	231
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/clone="B430002H17"
	/clone_lib="RIKEN full-length enriched, 4 days neonate male adipose"
	/sex="male"
	/tissue.type="adipose"
	/dev_stage="4 days neonate"
	/lab_host="DH10B"
	/note="Site_1: Sali; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken. Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'

FEATURES
source

```

BASE COUNT      34 a    61 c    98 g    38 t
ORIGIN
Query Match      84.0%; Score 16.8; DB 9; Length 231;
Best Local Similarity 90.0%; Pred. No. 5.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Caps 0;

QY      1 ggggtcacccggtgagggggg 20
          ||||| | |||||
Db      56 GGGGTACCCGGAGGGGG 75

Search completed: August 10, 2002, 02:11:17
Job time: 13138 sec

```

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki,
N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.,
Thermostabilization and thermoactivation of thermostable enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Kitsumai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,
Y. and Hayashizaki, Y.,
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.,
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 231)
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadota,K., Kawaga,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.,
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
Ono,F., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
Y., Shigenoto,Y., Shingawa,A., Shiraki,A., Sogabe,Y., Sugahara,Y.,
Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamana,I.,
Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al.)
Unpublished (2000).

COMMENT

COMMENT

COMMENT

FEATURES

source

**BASE COUNT
ORIGIN**

Query Match 84.0%; Score 16.8; DB 9; Length 231;
Best Local Similarity 90.0%; Pred. No. 5.1e+03;

```

FEATURES
source
Location/Qualifiers
1. .208
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_id="B430002L16"
/clone_lib="RIKEN full-length enriched, 4 days neonate
male adipose"
/sex="male"
/tissue_type="adipose"
/dev_stage="4 days neonate"
/lab_host="DH10B"
/note="Site_1: Sali; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken

```

KIKEN: Division of Experimental Animal Research, National Institute of Advanced Industrial Science and Technology, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8565, Japan

transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 229.0. Second

sequence [5' GAGAGAGAGATTCTCGAGTTAAATTAATCCCGCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I.¹⁰

```

Query Match      84.0%; Score 16.8; DB 9; Length 208;
Best Local Similarity 90.0%; Pred. No. 5e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ggggtccaccggtgagggggg 20
         |||||
Db       74  GGGGTCAACCGCGAGGGGG 93

```

RESULT	13
BB018784/c	
LOCUS	
DEFINITION	linear EST 23-JUN-2000
	230 bp mRNA
	BB018784 RIKEN full-length enriched, adult male testis (DH10B) Mus musculus cDNA clone 4930579N23 3', similar to D87896 Mus musculus phgpX mRNA for phospholipid hydroperoxide glutathione peroxidase, mRNA sequence.

VERSION	BB018784.1	GT:8190478
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
PREFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	

TITLE
JOURNAL
COMMENT

KONNO, H., AKIZAWA, H., AKASHI, R., AKIYAMA, Y., AKAKAWA, A., CARLIERI, P., ENDO, T., FUKUDA, S., FUKUNISHI, Y., HARA, A., HAYASU, N., HIROZANE, T., HORI, F., ISHII, Y., ISHIKAWA, J., ISHIKAWA, T., ITOH, M., IZAWA, M., KADOTA, K., KAGAWA, I., KAI, C., KAWAI, J., KIKUCHI, N., KIYOSAWA, H., KOJIMA, Y., KONDO, S., KOYA, S., KURIHARA, C., KUSAKABE, M., MATSUYAMA, T., MIKI, R., MIZUNO, Y., NAKAMURA, M., ODA, H., OKAZAKI, Y., ONO, T., OWA, C., SAITO, H., SAKAI, C., SATO, K., SHIBATA, K., SHIBATA, Y., SHIGEMOTO, Y., SHINGAWA, A., SHIRAKI, T., SOGABE, Y., SUGAHARA, Y., SUZUKI, H., SUZUKI, H., TAGAWA, A., TAKAHASHI, F., TOMINAGA, N., TOYA, T., TSUNODA, Y., WATAHIKI, A., WATANABE, S., YAMAMURA, T., YAMANAKA, I., YANO, R., YASUNISHI, A., YOKOTA, T., YOSHIDA, K., YOSHIKI, A., YOSHINO, M., MURAMATSU, M. and HAYASHIZAKI, Y.
RIKEN Mouse ESTs (Konno, H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

sequence [5' GAGAGAGATTCTCGAGTTAATAATTAATCCGCCGCCGCCGCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

BASE COUNT 69 a 86 c 56 g 60 t
ORIGIN

Query Match 85.0%; Score 17; DB 9; Length 271;

Best Local Similarity 100.0%; Pred. No. 4.3e+03; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 0;

QY 4 gtccacgggtgagggggg 20

Db 100 GTCACCGGTGAGGGGG 84

RESULT 8

BB593808

LOCUS BB593808 172 bp mRNA linear EST 30-NOV-2000
DEFINITION Mus musculus cDNA clone B430302P10 5' mRNA sequence.

ACCESSION BB593808.1

VERSION GI:11490410

KEYWORDS

SOURCE EST.

ORGANISM house mouse.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Aizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P.,

Hanagaki, T., Hayatsu, N., Hiraoka, T., Hirozane, T., Hodo, Y., Konno

Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J., Kojima, Y., Konno

Nomura, K., Numazaki, R., Okazaki, Y., Okido, T., Owa, C., Sakai, C.,

Sakai, K., Sasaki, D., Sato, K., Shibata, K., Shibata, Y., Shinagawa, A.,

Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka

T., Toya, T., Watahiki, A., Yamamura, T., Yasunishi, A., Yoshida, K.,

Yoshiki, A., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Aizawa, K. et al. 2000)

Unpublished (2000)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki

N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Thermotabilization and thermoactivation of thermostable enzymes by

trehalose and its application for the synthesis of full length

cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki

Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning; Methods Enzymol. 303,

19-44 (1999)

Please visit our web site (http://genome.rtc.riken.go.jp) for

further details.

Location/Qualifiers

1. .172

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="B430302P10"

/clone_lib="RIKEN full-length enriched, 4 days neonate

male adipose"

/sex="male"

FEATURES

source

/tissue_type="adipose"
/dev_stage="4 days neonate"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAAGGATCCAGAGCTCTTTTTTTTTTTTTTTN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 229.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGATTCTCGAGTTAATAATTAATCCGCCGCCGCCGCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

BASE COUNT 29 a 46 c 67 g 30 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 172;

Best Local Similarity 90.0%; Pred. No. 5e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggggtcaccgggtgagggggg 20

Db 49 GGGGTACCCCGGAGGGGG 68

RESULT 9

BB602331

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Aizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P.,

Hanagaki, T., Hayatsu, N., Hiraoka, T., Hirozane, T., Hodo, Y., Konno

Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J., Kojima, Y., Konno

H., Kusakabe, M., Matsuyama, T., Miyazaki, A., Nakamura, M., Nishi, K.,

Nomura, K., Numazaki, R., Okazaki, Y., Okido, T., Owa, C., Sakai, C.,

Sakai, K., Sasaki, D., Sato, K., Shibata, K., Shibata, Y., Shinagawa, A.,

Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka

T., Toya, T., Watahiki, A., Yamamura, T., Yasunishi, A., Yoshida, K.,

Yoshiki, A., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Aizawa, K. et al. 2000)

Unpublished (2000)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki

N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Thermotabilization and thermoactivation of thermostable enzymes by

trehalose and its application for the synthesis of full length

cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki

Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning; Methods Enzymol. 303,

19-44 (1999)

Please visit our web site (http://genome.rtc.riken.go.jp) for

further details.

Location/Qualifiers

1. .172

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="B430302P10"

/clone_lib="RIKEN full-length enriched, 4 days neonate

male adipose"

/sex="male"

Seq primer: -21M13 Forward
High quality sequence stop: 305
POLYA-Yes.

FEATURES

source

Location/Qualifiers

1. .305
/organism="Mus musculus"
/strain="129/Sv x 129/Sv-CP"
/db_xref="niaEST:C0317B07-3"
/db_xref="taxon:10090"
/clone="C0317B07"
/clone_lib="NIA Mouse Undifferentiated ES Cell cDNA Library (Long)"
/tissue_type="Undifferentiated ES Cell"
/cell_line="R1 ES cells"
/lab_host="DH10B"
/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were obtained from Dr. Kenneth R. Boheler (National Institute on Aging, USA). ES cells were cultured without feeder cells in the presence of LIF and BRL-conditioned media. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen].
5'-pGACTAGTTTATGACGGAGCGGCCCTTTT-3' from 14.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lona-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.4 kb. The library was constructed by Yulan Piao (NIA)."

75 a 71 c 100 g 59 t

BASE COUNT
ORIGIN

Query Match 87.0%; Score 17.4; DB 10; Length 305;
Best Local Similarity 94.7%; Pred. No. 3e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggtcaccggtgaggggg 19

||||| |||||||

Db 236 GGGTCACGGGTGAGGGG 254

RESULT 4

LOCUS

AV592822/c 466 bp mRNA linear EST 27-NOV-2001
DEFINITION AV592822 Bos taurus cartilage fetus Bos taurus cDNA clone
EICA003H03 3', mRNA sequence.

ACCESSION AV592822

VERSION AV592822.1 GI:9707979

KEYWORDS EST.

SOURCE cow.

ORGANISM

Bos taurus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE

AUTHORS

1 (bases 1 to 466)
Takasuga, A., Hirotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H.
and Sugimoto, Y.

TITLE

Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs

JOURNAL

Nucleic Acids Res. 29 (22), E108 (2001)

MEDLINE

COMMENT

21570554

Contact: Yoshikazu Sugimoto

Animal Genetics Division

Shirakawa Institute of Animal Genetics

Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan

Tel: 81-248-25-5641

Fax: 81-248-25-5725

Email: kazusugi@cocoa.ocn.ne.jp

Single pass sequencing.

This clone was obtained from a polyA-deleted cDNA library.

FEATURES

source

Location/Qualifiers

1. .466
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="EICA003H03"
/clone_lib="Bos taurus cartilage fetus"
/tissue_type="cartilage"
/dev_stage="fetus"
/lab_host="DH10B"
/note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"

95 a 124 c 170 g 74 t 3 others

BASE COUNT

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 466;

Best Local Similarity 94.7%; Pred. No. 3e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggtcaccggtgaggggg 19

||||| |||||||

Db 363 GGGTCCCCGTGAGGGG 345

RESULT 5

LOCUS

BE613321 961 bp mRNA linear EST 20-OCT-2000
DEFINITION 601452592T1 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3856198 3',
mRNA sequence.

ACCESSION BE613321

VERSION BE613321.1 GI:9894918

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 961)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-femail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM9584 row: h column: 23

High quality sequence start: 2

High quality sequence stop: 775.

FEATURES

source

Location/Qualifiers

1. .961
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3856198"
/clone_lib="NIH_MGC_66"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life

ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 1083;
 Best Local Similarity 95.0%; Pred. No. 1.3e+03;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggtcaccggtgagggggg 20
 ||||| ||||| ||||| |||||

Db 924 GGGGTCGCCGCGTGAGGGGG 905

RESULT 2
 BB714282/c standard; RNA; EST; 286 BP.

XX AC BB714282;
 SV BB714282.1

XX 09-OCT-2001 (Rel. 69, Created)
 DT 09-OCT-2001 (Rel. 69, Last updated, Version 1)

XX Mus musculus 9.5 days embryo parthenogenote cDNA, RIKEN full-length
 DE enriched library, clone: B130058N22, 3' end partial sequence.
 DE EST (expressed sequence tag).

XX Mus musculus (house mouse)
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

XX [1]
 RP 1-286
 RA Akimura T., Arakawa T., Carninci P., Furuno M., Hanagaki T., Hayatsu N.,
 RA Hiramoto K., Hiraoka T., Hirozane T., Imotani K., Ishii Y., Ito M.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Matsuyama T., Nakamura M.,
 RA Nishi K., Nomura K., Numasaki R., Okazaki Y., Okido T., Saito R., Sakai C.,
 RA Sakai K., Sakazume N., Sasaki D., Sato K., Shibata K., Shinagawa A.,
 RA Shiraki T., Sogabe Y., Suzuki H., Tagawa A., Takahashi F.,
 RA Takaku-Akahira S., Tanaka T., Tomaru A., Toya T., Watahiki A.,
 RA Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RT Submitted (28-SEP-2001) to the EMBL/GenBank/DBJ databases.
 RL Yoshihide Hayashizaki, The Institute of Physical and Chemical Research,
 RL (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic
 RL Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho,
 RL Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 RL (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/,
 RL Tel: 81-45-503-9222, Fax: 81-45-503-9216)

XX [2]
 RA Akimura T., Arakawa T., Carninci P., Furuno M., Hanagaki T., Hayatsu N.,
 RA Hiramoto K., Hiraoka T., Hirozane T., Imotani K., Ishii Y., Ito M.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Matsuyama T., Nakamura M.,
 RA Nishi K., Nomura K., Numasaki R., Okazaki Y., Okido T., Saito R., Sakai C.,
 RA Sakai K., Sakazume N., Sasaki D., Sato K., Shibata K., Shinagawa A.,
 RA Shiraki T., Sogabe Y., Suzuki H., Tagawa A., Takahashi F.,
 RA Takaku-Akahira S., Tanaka T., Tomaru A., Toya T., Watahiki A.,
 RA Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RT "RIKEN Encyclopedia of Mouse Full-length cDNAs";
 RL Unpublished.

XX [3]
 RA Konno H., Fukunishi Y., Shibata K., Itoh M., Carninci P., Sugahara Y.,
 RA Hayashizaki Y.;
 RT "Computer-based methods for the mouse full-length cDNA encyclopedia:
 RT real-time sequence clustering for construction of a nonredundant cDNA
 RT library";
 RL Genome Res. 11:281-289(2001).

XX [4]
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare
 RL full-length cDNA libraries for rapid discovery of new genes";
 XX Genome Res. 10:1617-1630(2000).

[5]
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M., Sumi N.,
 RA Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R.,
 RA Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaka S.,
 RA Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y.,
 RA Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y.,
 RA Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing
 RL pipeline with 384 multicapillary sequencer";
 XX Genome Res. 10:1757-1771(2000).

CC Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 CC further details.
 CC cDNA library was prepared and sequenced in Mouse Genome
 CC Encyclopedia Project of Genome Exploration Research Group in Riken
 CC Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 CC Division of Experimental Animal Research in Riken contributed to
 CC prepare mouse tissues.

XX Key Location/Qualifiers
 FH source 1..286
 FT /db_xref="taxon:10090"
 FT /sequenced_mol="cDNA to mRNA"
 FT /organism="Mus musculus"
 FT /clone="B130058N22"
 FT /clone_lib="RIKEN full-length enriched mouse cDNA library"
 FT /dev_stage="9.5 days embryo"
 FT /strain="C57BL/6J"
 FT /tissue_type="parthenogenote"

XX Sequence 286 BP; 70 A; 100 C; 57 G; 59 T; 0 other;
 SQ Query Match 90.0%; Score 18; DB 4; Length 286;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ggtcaccggtgagggggg 20
 ||||| ||||| ||||| |||||

Db 116 GGTACCGGTGAGGGGGG 99

RESULT 3
 BM195805
 LOCUS C0317B07-3 NIA Mouse Undifferentiated ES Cell cDNA Library (Long)
 DEFINITION Mus musculus cDNA clone C0317B07 3', mRNA sequence.
 ACCESSION BM195805
 VERSION BM195805.1 GI:17747413
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 305)
 Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Luo, A.,
 Jaradat, S.A., Boheler, K.R. and Ko, M.S.H.
 Systematic Analyses of NIA Mouse Undifferentiated ES Cell cDNA
 Library (Long)
 Unpublished (2001)
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: C0317 row: B column: 07

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:11:14 ; Search time 9068.22 seconds
(without alignments)
29.768 Million cell updates/sec

Title: US-09-672-126-24

Perfect score: 20
Sequence: 1 ggggtcacccgtgaggggg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vit:*

Pred. 30. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.4	92.0	1083	BF535962	BF535962 602051583
C 2	18.4	90.0	286	BD714282	BD714282 Mus muscu
C 3	17.4	87.0	305	BM195805	BM195805 C0317807
C 4	17.4	87.0	466	AV592822	AV592822 AV592822
C 5	17.4	87.0	961	BE613321	BE613321 601452592
C 6	17.4	87.0	1226	BG824786	BG824786 602728777
C 7	17.4	85.0	271	BB196041	BB196041 BB196041
C 8	16.8	84.0	172	BB593808	BB593808 BB593808
C 9	16.8	84.0	189	BB602331	BB602331 BB602331
C 10	16.8	84.0	191	AA390163	AA390163 mr37a03.r
C 11	16.8	84.0	196	AA427026	AA427026 ve82b07.r
C 12	16.8	84.0	208	BB593445	BB593445 BB593445
C 13	16.8	84.0	230	BB018784	BB018784 BB018784
C 14	16.8	84.0	231	BB451497	BB451497 BB451497
C 15	16.8	84.0	231	BB593432	BB593432 BB593432
C 16	16.8	84.0	239	W15812	W15812 mb51d07.r1
C 17	16.8	84.0	241	AV287998	AV287998 AV287998

C 18	16.8	84.0	248	9	AA823351	AA823351
C 19	16.8	84.0	252	9	AV106362	AV106362
C 20	16.8	84.0	256	9	AV212696	AV212696
C 21	16.8	84.0	261	9	AV123761	AV123761
C 22	16.8	84.0	281	9	BB720853	BB720853
C 23	16.8	84.0	286	9	AA198615	AA198615
C 24	16.8	84.0	343	9	AA183598	AA183598
C 25	16.8	84.0	347	9	AA795055	AA795055
C 26	16.8	84.0	347	10	H65336	H65336
C 27	16.8	84.0	359	9	AI641880	AI641880
C 28	16.8	84.0	364	9	AA791765	AA791765
C 29	16.8	84.0	365	10	BG662242	BG662242
C 30	16.8	84.0	368	10	W99168	W99168
C 31	16.8	84.0	372	9	AW164125	AW164125
C 32	16.8	84.0	393	9	AA087855	AA087855
C 33	16.8	84.0	394	9	AA871029	AA871029
C 34	16.8	84.0	403	9	AI386328	AI386328
C 35	16.8	84.0	408	9	AA681474	AA681474
C 36	16.8	84.0	418	9	AA058713	AA058713
C 37	16.8	84.0	423	9	AA985697	AA985697
C 38	16.8	84.0	423	9	AA433773	AA433773
C 39	16.8	84.0	431	10	W17480	W17480
C 40	16.8	84.0	431	12	AQ334452	AQ334452
C 41	16.8	84.0	434	10	BG662142	BG662142
C 42	16.8	84.0	450	9	AA596816	AA596816
C 43	16.8	84.0	455	9	AA003326	AA003326
C 44	16.8	84.0	456	9	AA562507	AA562507
C 45	16.8	84.0	458	9	AA239697	AA239697

ALIGNMENTS

RESULT 1
BF535962/c
LOCUS 602051583P1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4190612 5', linear EST 11-DEC-2000
DEFINITION mRNA sequence.
ACCESSION BF535962
VERSION BF535962.1 GI:11623330
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1083)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9518 row: f column: 21
High quality sequence stop: 658.
Location/Qualifiers
1. 1083
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4190612"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:
NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 264 a 344 c 263 g 211 t

THIS PAGE BLANK (USPTO)

KW Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
 KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
 XX Synthetic.
 XX
 XX
 XX
 XX Key Location/Qualifiers
 PH modified_base 1..20
 FT /*tag= a
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 XX
 XX WO200122990-A2.
 XX
 XX 05-APR-2001.
 XX
 XX 27-SEP-2000; 2000WO-US26527.
 XX
 XX 27-SEP-1999; 99US-0156147.
 XX
 XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX
 XX Hartmann G, Bratzler RL, Krieg A;
 XX WPI; 2001-290487/30.
 XX
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 XX
 XX Claim 201; Page 103; 168pp; English.
 XX
 XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX
 XX Sequence 20 BP; 3 A; 2 C; 12 G; 3 T; 0 other;
 SQ

Query Match 84.0%; Score 16.8; DB 22; Length 20;
 Best Local Similarity 90.0%; Pred. No. 2.1e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ggggtcacccggtgagggggg 20
 ||||| || |||||
 Db 1 ggggtcacccggtgagggggg 20
 ||||| || |||||

RESULT 15
 AAF98854
 ID AAF98854 standard; DNA; 20 BP.
 XX
 XX AAF98854;
 AC
 XX
 XX 11-JUN-2001 (first entry)
 XX
 XX Poly-G immunostimulatory nucleic acid SEQ ID NO: 135.
 DE
 DE Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
 KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
 XX Synthetic.
 OS
 XX WO200122990-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 27-SEP-2000; 2000WO-US26527.
 PF

XX
 PR 27-SEP-1999; 99US-0156147.
 XX
 XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX
 XX Hartmann G, Bratzler RL, Krieg A;
 XX WPI; 2001-290487/30.
 XX
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 XX
 XX Disclosure; Page 24; 168pp; English.
 XX
 XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX
 XX Sequence 20 BP; 3 A; 2 C; 12 G; 3 T; 0 other;
 SQ

Query Match 84.0%; Score 16.8; DB 22; Length 20;
 Best Local Similarity 90.0%; Pred. No. 2.1e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ggggtcacccggtgagggggg 20
 ||||| || |||||
 Db 1 ggggtcacccggtgagggggg 20
 ||||| || |||||

Search completed: August 10, 2002, 03:21:51
 Job time: 13682 sec

```
XX CpG motif containing oligonucleotide SEQ ID #5.
DE Immune system stimulator; CpG motif; CpG receptor; CpG-R; antibacterial;
XX immune response; vaccine adjuvant; tumour immunotherapy; allergy;
KW anti-inflammatory; cystic fibrosis; sepsis; heart disease; chlamydia;
KW inflammatory bowel disease; arthritis; multiple sclerosis; ss.
XX Unidentified.
XX
XX Key Location/Qualifiers
FH modified_base 1..20
FT /*tag= a
FT /mod_base= OTHER
FT /note= "Phosphorothioate internucleoside linkages"
XX
XX WO200132877-A2.
XX
XX 10-MAY-2001.
XX
XX 01-NOV-2000; 2000WO-US41735.
XX
XX 02-NOV-1999; 99US-0163157.
XX 24-NOV-1999; 99US-0167389.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Mackichan ML;
XX
XX WPI; 2001-343486/36.
XX
XX Novel CpG receptor and nucleic acid molecule encoding the receptor, for
XX modulating immune response and for identifying compounds of therapeutic
XX use which bind and/or modulate the activity of the receptor
XX
XX Example 1; Page 14; 41pp; English.
XX
XX Unmethylated CG dinucleotide sequences are commonly found in bacterial
XX DNA, and have been found to stimulate the innate immune system. Natural
XX killer and T cells are activated by exposure to oligonucleotides
XX containing CpG motifs. Oligonucleotides containing CpG motifs can be used
XX as adjuvants in vaccines. The present invention relates to a CpG
XX receptor. The CpG receptor contains a Toll homology domain (THD). The
XX Toll receptor family are associated with responses to pathogens. CpG
XX oligonucleotides may act as stimulators of various immune responses. The
XX CpG receptor or cells expressing the receptor are useful for identifying
XX a compound which binds to or modulates an activity of the CpG receptor.
XX The compounds are useful in e.g. vaccine adjuvants promoting
XX cell-mediated immune responses, antibacterials, (e.g. protection from
XX Listeria infection), tumour immunotherapy, allergy treatment, (e.g.
XX suppressing IgE in human PBMC, shifting from Th2 to Th1) and as
XX anti-inflammatory agents (e.g. for use in cystic fibrosis, sepsis, heart
XX disease, chlamydia, inflammatory bowel disease, arthritis and multiple
XX sclerosis). The present sequence represents a CpG motif containing
XX oligonucleotide used in examples demonstrating that CpG oligonucleotides
XX can activate the MAPK pathways and NF-kappaB.
XX
XX Sequence 20 BP; 3 A; 2 C; 12 G; 3 T; 0 other;
XX
XX Query Match 84.0%; Score 16.8; DB 22; Length 20;
XX Best Local Similarity 90.0%; Pred. NO. 2.1e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 ggggtcacccggtgagggggg 20
XX ||||| |||||
XX Db 1 ggggtcacccggtgagggggg 20
XX
XX RESULT 13
XX AAF98731
XX ID AAF98731 standard; DNA; 20 BP.
XX
XX
```

```
AC AAF98731;
XX
XX 11-JUN-2001 (first entry)
XX
XX Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 1.
DE
XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
XX viral infection; phosphorothioate backbone; palindrome; cancer; ds.
KW
XX Synthetic.
XX
XX Key Location/Qualifiers
FH modified_base 1..2
FT /*tag= a
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
XX
XX modified_base 15..19
FT /*tag= b
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
XX
XX WO200122990-A2.
XX
XX 05-APR-2001.
XX
XX 27-SEP-2000; 2000WO-US26527.
XX
XX 27-SEP-1999; 99US-0156147.
XX
XX (COLE-) COLEY PHARM GROUP INC.
XX (IOWA ) UNIV IOWA RES FOUND.
XX
XX Hartmann G, Bratzler RL, Krieg A;
XX
XX WPI; 2001-290487/30.
XX
XX Improving the efficacy of treatments involving the administration of
XX interferon-alpha by co-administering an isolated immunostimulatory
XX nucleic acid
XX
XX Claim 19; Page 73; 168pp; English.
XX
XX The present invention describes an improvement to a method requiring the
XX administration of interferon alpha (IFN-alpha), involving administering
XX an immunostimulatory nucleic acid (ISNA). The sequences of a number of
XX such nucleic acids are also provided. These may comprise oligonucleotides
XX with phosphorothioate backbones, palindromes, or G-rich sequences. The
XX sequences of the invention are useful in the treatment of proliferative
XX diseases, such as cancers, and viral infections. The present sequence is
XX an example of an immunostimulatory oligonucleotide.
XX
XX Sequence 20 BP; 3 A; 2 C; 12 G; 3 T; 0 other;
XX
XX Query Match 84.0%; Score 16.8; DB 22; Length 20;
XX Best Local Similarity 90.0%; Pred. NO. 2.1e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 ggggtcacccggtgagggggg 20
XX ||||| |||||
XX Db 1 ggggtcacccggtgagggggg 20
XX
XX RESULT 14
XX AAF98736
XX ID AAF98736 standard; DNA; 20 BP.
XX
XX AAF98736;
XX
XX 11-JUN-2001 (first entry)
XX
XX Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 6.
XX
```

KW therapeutic; allergy; asthma; cancer; autoimmune disorder; infection;
 KW bio-warfare; vaccine; antisense therapy; eczema; allergic rhinitis;
 KW coryza; hay fever; urticaria; hives; food allergy; atopic condition;
 KW hepatitis; human immunodeficiency virus; HIV; malaria; Francisella;
 KW lupus erythematosus; rheumatoid arthritis; multiple sclerosis;
 KW schistosomiasis; tuberculosis; acquired immunodeficiency syndrome; AIDS;
 KW Leishmania; Ebola; Anthrax; Listeria; ss.

Synthetic.

WO200151500-A1.

19-JUL-2001

12-JAN-2001: 2001WO-IIS01122

14-JAN-2000: 2000HS-0176115

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Klinman D, Ishii K, Verthelvy D:

WPI: 2001-442129/47

Oligodeoxynucleotides for inducing an immune response to treat and prevent an allergic reaction, cancer, an autoimmune disorder and symptoms resulting from exposure to bio-warfare agents, comprise multiple CpG sequences -

Claim 5; Page 42; 48pp; English.

AAS09551-AAS093662 represent oligodeoxynucleotides (ODN) of at least 10 nucleotides comprising multiple CpG sequences, where one of the CpG sequences is different from another of the multiple CpG sequences. The ODN are useful for inducing an immune response, preferably a cell-mediated immune response, involving non-B cell activation, interferon gamma (IFN-gamma) production or a humoral immune response involving B cell activation, antibody and interleukin-6 production in a host, for treating, preventing or ameliorating an allergic reaction, e.g. asthma, cancer, e.g. solid tumour cancer, a disease associated with the immune system e.g. autoimmune disorder or an immune system deficiency, infection or a symptom resulting from exposure to bio-warfare agent in a human. The induction of immune response improves the efficacy of a vaccine and is used in antitense therapy. The ODN are useful for treating, preventing, or ameliorating allergic reactions, including eczema, allergic rhinitis or coryza, hay fever, bronchial asthma, urticaria (hives), food allergies and other atopic conditions, for improving the efficacy of vaccines against hepatitis A, B and C, human immunodeficiency virus (HIV) and malaria, for treating immune system deficiencies, e.g. lupus erythematosus and autoimmune diseases such as rheumatoid arthritis and multiple sclerosis, infections including Francisella, schistosomiasis, tuberculosis, acquired immunodeficiency syndrome (AIDS), Leishmania and symptoms resulting from exposure of bio-warfare agent, including Ebola, Anthrax and Listeria.

Sequence 20 BP; 3 A; 2 C; 12 G; 3 T; 0 other:

```
Query Match      84.0%; Score 16.8; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy 1 ggggtcaccggtgaggggg 20
 ||||| || |||||
 Db 1 ggggtcaacqttgaqqqqq 20

RESULT 11

RESUL II
AAH50658
ID AAH50658 standard: DNA: 20 BP.

XX
AC AAH50658;

XX

22-AUG-2001 (first entry)

Immune response modulating related oligonucleotide SEQ ID NO: 90.

KW Immunostimulatory; inducing; natural killer cell; lytic activity;
 KW unmethylated CpG dinucleotide; immune response; B cell proliferation;
 KW Th1; immune activation; interleukin 6; IL-6; interferon gamma;
 KW IFN-gamma; cytokine; ss.

Synthetic.

PN US6239116-B1.

XX
PD
29-MAY-2001.

XX
PF 30-OCT-1997: 97US-0960774.

XX 30-OCT-1996; 96US-0738652.

XX (IOWA) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GROUP INC.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES

AA
PI
Krieg AM, Kline JN:

XX
DR WPT: 2001-380456/40.

Methods for inducing IL-6, interferon-gamma or IL-12, or stimulating natural killer cell lytic activity in a human, comprise administering to the subject or exposing a natural killer cell to immunostimulatory nucleic acids -

XX
PS
Disclosure: Column 91: 74pp: English:

The present invention describes methods for inducing interleukin 6 (IL-6), interferon-gamma (IFN-gamma) or IL-12, or for stimulating natural killer cell lytic activity. The methods comprise administering to the subject or exposing a natural killer cell to an immunostimulatory nucleic acid. Also described are: (1) inducing IL-6 in a subject comprising administering to the subject to induce IL-6 in the subject the immunostimulatory nucleic acid; (2) stimulating natural killer cell lytic activity comprising exposing a natural killer cell to the immunostimulatory nucleic acid to stimulate natural killer cell lytic activity; (3) inducing interferon-gamma in a subject to treat an immune system deficiency comprising administering to the subject to induce interferon-gamma production, the immunostimulatory nucleic acid; and (4) inducing IL-12 in a subject comprising administering to the subject the immunostimulatory nucleic acid. The methods are useful for inducing IL-6, interferon-gamma or IL-12, or stimulating natural killer cell lytic activity in a subject, particularly a human. The methods are particularly useful for modulating an immune response. AAH50571 to AAH50671 represent oligonucleotide sequences used in the exemplification of the present invention.

XX
SO
Sequence 20 BP: 3 A: 2 C: 12 G: 3 T: 0 other:

```
Query Match      84.0%; Score 16.8; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy 1 ggggtcaccggtgagggggg 20
||||| ||| |||||

Dp 1 ggggtcaacgttgagggggg 20

RESULT 12

RESULT 12
AAH20394
ID AAH20394 standard: DNA: 20 BP.

XX
AC AAH20394:XX
DT 03-AUG-2001 (first entry)

```

XX OS Synthetic.
XX PN WO9852581-A1.
XX PD 26-NOV-1998.
XX PF 20-MAY-1998; 98WO-US10408.
XX PR 20-MAY-1997; 97US-0047233.
XX PR 20-MAY-1997; 97US-0047209.
XX PA (OTTA-) OTTAWA CIVIC HOSPITAL LOEB RES INST.
XX PA (QIAG-) QIAGEN GMBH.
XX PA (IOWA-) UNIV IOWA RES FOUND.
XX PI Davis HL, Kriegl AM, Schorr J, Wu T;
XX WPI; 1999-059712/05.
XX PT Use of neutralising CpG and stimulating CpG motifs in DNA vectors -
XX PT for enhancing the immunostimulatory effect of an antigen or
XX PT enhancing the expression of a therapeutic polypeptide
XX PS Example 1; Page 64; 109pp; English.
XX CC AAV74237-V74253 are oligodeoxynucleotide (ODN) primers used to describe
XX CC a method for enhancing the immunostimulatory effect of an antigen
XX CC encoded by nucleic acid contained in a nucleic acid construct. The
XX CC method involves determining the CpG-N and CpG-S motifs present in the
XX CC construct, removing neutralising CpG (CpG-N) motifs and optionally
XX CC inserting stimulatory CpG (CpG-S) motifs in the construct, thereby
XX CC producing a nucleic acid construct having enhanced immunostimulatory
XX CC efficacy. The method can be used for immunisation against viral antigens,
XX CC e.g. from hepatitis B virus (HBV), bacterial antigens or an antigen
XX CC derived from a parasite. They can also be used for expression of a
XX CC therapeutic polypeptide, e.g. growth factors, toxins, tumour suppressors,
XX CC cytokines, apoptotic proteins, interferons, hormones, clotting factors,
XX CC ligands and receptors.
XX SQ Sequence 20 BP; 3 A; 2 C; 12 G; 3 T; 0 other;

Query Match 84.0%; Score 16.8; DB 20; Length 20;
Best Local Similarity 90.0%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ggggtcacccgtgagggggg 20
||||| || |||||
Db 1 ggggtcacccgtgagggggg 20

RESULT 9
AAAA0449
ID AAA90449 standard; DNA; 20 BP.
XX AA90449;
XX AC AA90449;
XX AC AA90449;
XX DT 10-JAN-2001 (first entry)
XX DE CpG adjuvant oligonucleotide, SEQ ID NO:3.
XX KW CpG oligonucleotide: CpG motif; adjuvant; microdroplet emulsion;
XX KW microemulsion; adsorbent microparticle; vaccine; Th1 immune response;
XX KW viral infection; bacterial infection; parasitic infection; HCV; HBV;
XX KW hepatitis C virus; hepatitis B virus; herpes simplex virus; HSV; HIV;
XX KW human immunodeficiency virus; cytomegalovirus; CMV; influenza virus;
XX KW rabies virus; cholera; diphtheria; tetanus; pertussis;
XX KW Helicobacter pylori; Haemophilus influenzae; malaria; ss.
XX OS Synthetic.
XX PN WO200050006-A2.

```

```

XX PD 31-AUG-2000.
XX PF 09-FEB-2000; 200WO-US03331.
XX PR 26-FEB-1999; 99US-0121858.
XX PR 29-JUL-1999; 99US-0146391.
XX PR 28-OCT-1999; 99US-0161997.
XX PA (CHIR ) CHIRON CORP.
XX PI O'Hagan D, Ott GS, Donnelly J, Kazzaz J, Uguzzoli M, Singh M;
XX PI Barackman J;
XX DR WPI; 2000-587123/55.
XX PT Microemulsion having an adsorbent surface comprising a microdroplet
XX PT emulsion consisting of a metabolizable oil and an emulsifying agent
XX PT which is a detergent, useful as a vaccine to treat bacterial, viral,
XX PT and parasitic infection
XX PS Claim 17; Page 40; 95pp; English.
XX CC The invention relates to a microdroplet emulsion (microemulsion) with an
XX CC adsorbent surface, and which comprises a metabolisable oil and an
XX CC emulsifying agent (a detergent). It also relates to a composition
XX CC comprising the microemulsion and a microparticle with an adsorbent
XX CC surface, where the microparticle comprises a polymer selected from a
XX CC poly(alpha-hydroxy acid), a polyhydroxy butyric acid, a
XX CC polycaprolactone, a polyorthoester, a polyether, a polyether, and a
XX CC polycyanacrylate, and a second detergent. The surface of the
XX CC microparticles efficiently adsorb biologically active macromolecules such
XX CC as DNA, polypeptides, antigens, hormones, pharmaceuticals, enzymes,
XX CC mediators of transcription or translation, metabolic intermediates and
XX CC adjuvants. Additionally, a second biologically active molecule may be
XX CC encapsulated within the microparticle. The microemulsion can be used in
XX CC methods of immunising a host animal, particularly a human, against a
XX CC viral, bacterial or parasitic infection, and in methods of increasing a
XX CC Th1 immune response. The microemulsions (having the appropriate antigens
XX CC adsorbed) may be particularly used as vaccines for hepatitis C virus
XX CC (HCV), hepatitis B virus (HBV), herpes simplex virus (HSV), human
XX CC immunodeficiency virus (HIV), cytomegalovirus (CMV), influenza virus, and
XX CC rabies virus; the bacteria which cause cholera, diphtheria, tetanus and
XX CC pertussis; Helicobacter pylori and Haemophilus influenzae; and
XX CC malaria-causing parasites. Sequences AAA90447-A90467 represent Th1
XX CC lymphocyte stimulating oligonucleotides containing at least one CpG motif
XX CC which are claimed for use as adjuvants in the compositions of the
XX CC invention.
XX SQ Sequence 20 BP; 3 A; 2 C; 12 G; 3 T; 0 other;

Query Match 84.0%; Score 16.8; DB 21; Length 20;
Best Local Similarity 90.0%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ggggtcacccgtgagggggg 20
||||| || |||||
Db 1 ggggtcacccgtgagggggg 20

RESULT 10
AAS09639
ID AAS09639 standard; DNA; 20 BP.
XX AA09639;
XX AC AAS09639;
XX DT 26-SEP-2001 (first entry)
XX DE Immunoreactive CpG sequence-containing oligonucleotide #89.
XX KW CpG sequence: immune response; non-B cell activation; interferon gamma;
XX KW IFN-gamma; humoral; antibody production; interleukin-6 production;

```

XX AAV27654;
 XX 01-OCT-1998 (first entry)
 XX Immunostimulatory oligodeoxyribonucleotide of the invention.
 XX Immunostimulatory; oligodeoxyribonucleotide; ODN;
 XX unmethylated CpG dinucleotide; activate; lymphocyte; immune response;
 XX Th2; cytokine; treatment; prevention; asthma; autoimmune disease;
 XX desensitisation therapy; artificial adjuvant; antibody generation; ss.
 XX Synthetic.
 XX WO9818810-A1.
 XX 07-MAY-1998.
 XX 30-OCT-1997; 97WO-US19791.
 XX 30-OCT-1996; 96US-0738652.
 XX (IOWA) UNIV IOWA RES FOUND.
 XX Kline JN, Krieg AM;
 XX WPI; 1998-272127/24.
 XX New immunostimulatory nucleic acid molecules - which contain at
 XX least one unmethylated CpG dinucleotide, used for treating e.g.
 XX tumours, infections or autoimmune disease
 XX Claim 26; Page 83; 109pp; English.
 XX AAV27641-751 represent immunostimulatory oligodeoxyribonucleotides
 XX (ODNs) of the invention. The ODNs contain at least one unmethylated CpG
 XX dinucleotide, and have the formula:
 XX 5' N1X1CGX2N2 3', where at least one nucleotide separates consecutive
 XX CpGs, X1 is adenine, guanine, or thymine, X2 is cytosine or thymine, N
 XX is any nucleotide and N1+N2 is 0-26 bases with the provision that N1 and
 XX N2 does not contain a CCGG tetramer or more than one CCG or CGG trimer.
 XX OR 5' NX1X2CGX3X4N 3', where at least one nucleotide separates
 XX consecutive CpGs, X1 and X2 are selected from GpT, GpG, GpA, ApT and ApA,
 XX X3 and X4 are selected from Tpt or Cpt, N is any nucleotide and N1+N2 is
 XX 0-26 bases with the provision that N1 and N2 does not contain a CCGG
 XX tetramer or more than one CCG or CGG trimer.
 XX The ODNs activate lymphocytes in a subject and redirect a subject's
 XX immune response from a Th2 to a Th1 (e.g. by inducing monocyte cells
 XX and other cells to produce Th1 cytokines, including IL-12, IFN-gamma and
 XX GM-CSF). The ODNs can be used to treat or prevent an asthmatic disorder,
 XX autoimmune diseases, in desensitisation therapy, as an artificial
 XX adjuvant during antibody generation in a mammal such as a mouse or a
 XX human.
 XX Sequence 20 BP; 3 A; 2 C; 12 G; 3 T; 0 other;
 XX
 XX Query Match 84.0%; Score 16.8; DB 19; Length 20;
 XX Best Local Similarity 90.0%; Pred. No. 2.1e+02;
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ggggtcacccgttgagggggg 20
 Db 1 ggggtcacccgttgagggggg 20
 RESULT 7
 AAV74238
 ID AAV74238 standard; DNA; 20 BP.
 XX
 XX AAV74238;
 XX
 XX 15-MAR-1999 (first entry)
 XX
 XX CpG-N motif; immunostimulation; antigen; CpG-S motif; immunisation; ODN;
 XX viral antigen; bacterial antigen; parasite; therapeutic; growth factor;
 XX toxin; tumour suppressor; cytokine; apoptotic protein; interferon;
 XX hormone; clotting factor; ligand; receptor; oligodeoxynucleotide; ss.

XX CpG-N motif S-ODN 1628 DNA.
 XX
 XX CpG-N motif; immunostimulation; antigen; CpG-S motif; immunisation; ODN;
 XX viral antigen; bacterial antigen; parasite; therapeutic; growth factor;
 XX toxin; tumour suppressor; cytokine; apoptotic protein; interferon;
 XX hormone; clotting factor; ligand; receptor; oligodeoxynucleotide; ss.
 XX Synthetic.
 XX WO9852581-A1.
 XX 26-NOV-1998.
 XX 20-MAY-1998; 98WO-US10408.
 XX 20-MAY-1997; 97US-0047233.
 XX 20-MAY-1997; 97US-0047209.
 XX (OTTA-) OTTAWA CIVIC HOSPITAL LOEB RES INST.
 XX (QIAG-) QIAGEN GMBH.
 XX (IOWA-) UNIV IOWA RES FOUND.
 XX Davis HL, Krieg AM, Schorr J, Wu T;
 XX WPI; 1999-059712/05.
 XX Use of neutralising CpG and stimulating CpG motifs in DNA vectors -
 XX for enhancing the immunostimulatory effect of an antigen or
 XX enhancing the expression of a therapeutic polypeptide
 XX Example 1; Page 64; 109pp; English.
 XX AAV74237-V74253 are oligodeoxynucleotide (ODN) primers used to describe
 XX a method for enhancing the immunostimulatory effect of an antigen
 XX encoded by nucleic acid contained in a nucleic acid construct. The
 XX method involves determining the CpG-N and CpG-S motifs present in the
 XX construct, removing neutralising CpG (CpG-N) motifs and optionally
 XX inserting stimulatory CpG (CpG-S) motifs in the construct, thereby
 XX producing a nucleic acid construct having enhanced immunostimulatory
 XX efficacy. The method can be used for immunisation against viral antigens,
 XX e.g. from hepatitis B virus (HBV), bacterial antigens or an antigen
 XX derived from a parasite. They can also be used for expression of a
 XX therapeutic polypeptide, e.g. growth factors, toxins, tumour suppressors,
 XX cytokines, apoptotic proteins, interferons, hormones, clotting factors,
 XX ligands and receptors.
 XX Sequence 20 BP; 3 A; 2 C; 12 G; 3 T; 0 other;
 XX
 XX Query Match 84.0%; Score 16.8; DB 20; Length 20;
 XX Best Local Similarity 90.0%; Pred. No. 2.1e+02;
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ggggtcacccgttgagggggg 20
 Db 1 ggggtcacccgttgagggggg 20
 RESULT 8
 AAV74245
 ID AAV74245 standard; DNA; 20 BP.
 XX
 XX AAV74245;
 XX
 XX 15-MAR-1999 (first entry)
 XX
 XX CpG-N motif SOS-ODN 1585 DNA.
 XX CpG-N motif; immunostimulation; antigen; CpG-S motif; immunisation; ODN;
 XX viral antigen; bacterial antigen; parasite; therapeutic; growth factor;
 XX toxin; tumour suppressor; cytokine; apoptotic protein; interferon;
 XX hormone; clotting factor; ligand; receptor; oligodeoxynucleotide; ss.

CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 XX
 SQ Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 8.5;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcacccgtgagggggg 20

|||||
 Db 1 ggggtcacccgtgagggggg 20

RESULT 4

AAT16894

ID AAT16894 standard; DNA; 20 BP.

AC AAT16894;

DT 06-SEP-1996 (first entry)

DE Immunomodulatory oligonucleotide contg. unmethylated C-G dinucleotide.

KW Unmethylated; immunomodulator; B cell activation; vaccine;

KW response stimulation; autoimmune disease; infection; ss.

OS Synthetic.

XX WO9602555-A1.

PD 01-FEB-1996.

PF 07-FEB-1995; 95WO-US01570.

PR 15-JUL-1994; 94US-0276358.

PA (IOWA) UNIV IOWA STATE RES FOUND INC.

XX Krieg AM;

DR WPI; 1996-105847/11.

XX Immunomodulatory oligo:nucleotide(s) contg. an un-methylated CpG
 PT di-nucleotide - used for stimulating activity or when methylated
 PT for inhibitory activity

XX Claim 5; Page 39; 45pp; English.

XX AAT16894-rl6898 are immunomodulatory oligonucleotides contg. at least
 CC one unmethylated C-G dinucleotide. The oligonucleotides can be used
 CC to activate B cells and natural killer cells. They can be used for
 CC treating, preventing or ameliorating an immune system deficiency,
 CC e.g. a tumour, cancer or a viral, fungal, bacterial or parasitic
 CC infection. They are also useful in stimulating a subject's response
 CC to a vaccine.

XX Sequence 20 BP; 3 A; 2 C; 12 G; 3 T; 0 other;

Query Match

Best Local Similarity 84.0%; Score 16.8; DB 17; Length 20;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggtcacccgtgagggggg 20

|||||

Db 1 ggggtcacccgtgagggggg 20

RESULT 5

AAV47684

ID AAV47684 standard; DNA; 20 BP.

AC AAV47684;

XX 20-NOV-1998 (first entry)

XX Unmethylated CpG dinucleotide 1585.

XX Unmethylated CpG dinucleotide; immune response; bacterial meningitis;
 KW natural killer cell activation; NK cell; Th2 response; neonatal sepsis;
 KW pulmonary disorder; asthma; environmentally induced airway disease;
 KW bacterial infection; endotoxaemia; therapy; cystic fibrosis;
 KW inflammatory bowel disease; ss.

OS Synthetic.

XX WO9837919-A1.

PD 03-SEP-1998.

PF 25-FEB-1998; 98WO-US03678.

PR 28-FEB-1997; 97US-0039405.

XX (IOWA) UNIV IOWA RES FOUND.

XX Krieg AM, Schwartz DA;

XX WPI; 1998-480941/41.

XX Use of nucleic acids containing an unmethylated CpG - for treating a
 PT subject having or at risk of having an acute decrement in air flow
 PT or inhibiting an inflammatory response

XX Claim 35; Page 27; 65pp; English.

XX This sequence represents an unmethylated CpG dinucleotide, and can be
 CC used in the method of the invention. The method is for treating a subject
 CC having, or at risk of having an acute decrement in air flow, comprising
 CC administering a nucleic acid sequence containing at least one
 CC unmethylated CpG. The nucleic acids containing an unmethylated CpG
 CC dinucleotide affect an immune response in a subject by activating natural
 CC killer cells (NK) or redirecting a subject's immune response from a Th2
 CC to a Th1 response by inducing monocytic and other cells to produce Th1
 CC cytokines. They can be used to treat pulmonary disorders having an
 CC immunologic component, such as asthma or environmentally induced airway
 CC disease. They can also be used to treat diseases associated with
 CC Gram-positive bacterial infections or endotoxaemia including bacterial
 CC meningitis, neonatal sepsis, cystic fibrosis, inflammatory bowel disease
 CC and liver cirrhosis, Gram-negative pneumonia, Gram-negative abdominal
 CC abscess, haemorrhagic shock, disseminated intravascular coagulation, or
 CC an inflammatory response to lipopolysaccharide.

XX Sequence 20 BP; 3 A; 2 C; 12 G; 3 T; 0 other;

Query Match

Best Local Similarity 84.0%; Score 16.8; DB 19; Length 20;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggtcacccgtgagggggg 20

|||||

Db 1 ggggtcacccgtgagggggg 20

RESULT 6

AAV27654

ID AAV27654 standard; DNA; 20 BP.

XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX Hartmann G, Bratzler RL, Krieg A;
 XX WPI; 2001-290487/30.
 DR
 XX
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 XX
 XX Claim 201; Page 103; 168pp; English.
 PS
 XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX
 XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;
 SQ

Query Match 100.0%; Score 20; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
 |||||
 Db 1 ggggtcacccggtgagggggg 20
 |||||

RESULT 2
 AAF99774
 ID AAF99774 standard; DNA; 20 BP.
 XX
 AC AAF99774;
 XX
 DT 12-JUN-2001 (first entry)
 XX
 DE Immunostimulatory nucleic acid #890.
 XX
 KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX
 OS Synthetic.
 XX
 PN WO200122972-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 25-SEP-2000; 2000WO-US26383.
 XX
 PR 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX
 PA (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 XX
 PI Krieg AM, Schetter C, Vollmer J;
 XX
 DR WPI; 2001-273485/28.
 XX
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 XX
 XX Claim 101; Page 57; 338pp; English.

XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: The present sequence may have a phosphorothioate backbone.
 XX
 XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;
 SQ

Query Match 100.0%; Score 20; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
 |||||
 Db 1 ggggtcacccggtgagggggg 20
 |||||

RESULT 3
 AAF99837
 ID AAF99837 standard; DNA; 20 BP.
 XX
 AC AAF99837;
 XX
 DT 12-JUN-2001 (first entry)
 XX
 DE Immunostimulatory nucleic acid #953.
 XX
 KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX
 OS Synthetic.
 XX
 PN WO200122972-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 25-SEP-2000; 2000WO-US26383.
 XX
 PR 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX
 PA (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 XX
 PI Krieg AM, Schetter C, Vollmer J;
 XX
 DR WPI; 2001-273485/28.
 XX
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 XX
 XX Claim 101; Page 58; 338pp; English.
 PS
 XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:21:50 ; Search time 1145.36 Seconds
(without alignments)
29,980 Million cell updates/sec

Title: US-09-672-126-24
Perfect score: 20
Sequence: 1 ggggtcacccggtgaggggg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Capext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	AAF98754	Human IFN-alpha im
2	20	100.0	20	AAF99774	Immunostimulatory
3	20	100.0	20	AAF99837	Immunostimulatory
4	16.8	84.0	20	AAT16894	Immunomodulatory o
5	16.8	84.0	20	AAV47684	Unmethylated CpG'd
6	16.8	84.0	20	AAV27654	Immunostimulatory
7	16.8	84.0	20	AAV74238	CpG-N motif S-ODN
8	16.8	84.0	20	AAV74245	CpG-N motif SOS-OD
9	16.8	84.0	20	AAA90449	CpG adjuvant oligo

10	16.8	84.0	20	22	AA509639	Immunoreactive CpG
11	16.8	84.0	20	22	AAH50658	Immune response mo
12	16.8	84.0	20	22	AAH20394	CpG motif containi
13	16.8	84.0	20	22	AAF98731	Human IFN-alpha im
14	16.8	84.0	20	22	AAF98736	Human IFN-alpha im
15	16.8	84.0	20	22	AAF98854	Poly-G Immunostimu
16	16.8	84.0	20	22	AAF99390	Immunostimulatory
17	16.8	84.0	20	22	AAF99567	Immunostimulatory
18	16.8	84.0	20	22	AAF99763	Immunostimulatory
19	16.8	84.0	20	22	AAF99764	Immunostimulatory
20	16.8	84.0	20	22	AAF99504	Immunostimulatory
21	16.8	84.0	20	22	AAF27750	P. falciparum vacc
22	16.8	84.0	20	22	AAAC80669	Immunogenic CpG ol
23	16.8	84.0	20	22	AAAC80669	CG motif and CFA C
24	16.8	84.0	20	22	AAAC92361	Oligonucleotide 15
25	16.8	84.0	20	22	AAH19262	Human IFN-alpha im
26	16.8	84.0	20	22	AAF98747	Immunostimulatory
27	16.8	84.0	20	22	AAF98875	Immunostimulatory
28	16.8	84.0	20	22	AAF99742	Immunostimulatory
29	16.8	84.0	20	22	AAF99797	Immunostimulatory
30	16.8	84.0	20	22	AAF99798	Immunostimulatory
31	16.8	84.0	20	22	AAF99389	Immunostimulatory
32	16.8	84.0	20	22	AAK70947	Human immune/haema
33	16.8	84.0	20	22	AAK70948	Human immune/haema
34	16.8	84.0	20	22	AAK57709	Human immune/haema
35	16.8	84.0	20	22	ABA21476	Human nervous syst
36	16.8	84.0	20	22	AAAD17185	Streptomyces nous
37	16.4	82.0	3435	17	AAAT55869	Streptomyces nous
38	16.4	82.0	3443	23	AAAT55869	Human DNA polymera
39	15.8	79.0	19	22	AAAT55869	DNA encoding novel
40	15.8	79.0	19	22	AAAT55869	Immunoreactive CpG
41	15.8	79.0	40	21	AAZ96149	Immunogenic CpG ol
42	15.8	79.0	795	21	AAAT55823	Polynucleotide seq
43	15.8	79.0	1700	16	AAAT55823	S. lavendulae mnci
44	15.8	79.0	1700	16	AAAT55823	Cepharmycin biosynt
45	15.8	79.0	1761	22	AAAI60814	cmch (ORF10) encod
						Human polynucleoti

ALIGNMENTS

RESULT	1
ID	AAF98754 standard; DNA; 20 BP.
XX	AAF98754;
AC	AAF98754;
XX	
DT	11-JUN-2001 (first entry)
XX	
DE	Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 24.
XX	
KW	Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW	viral infection; phosphorothioate backbone; palindrome; cancer; ds.
OS	Synthetic.
XX	
XX	
PH	Key Location/Qualifiers
FT	modified_base 1..2 a
FT	/*tag=
FT	/mod_base= "OTHER"
FT	/note= "phosphorothioate linkage"
FT	modified_base 15..19
FT	/*tag= b
FT	/mod_base= "OTHER"
FT	/note= "phosphorothioate linkage"
FT	
XX	
XX	
PN	WO200122990-A2.
XX	
XX	
PD	05-APR-2001.
XX	
XX	
PF	27-SEP-2000; 2000WO-US26527.
XX	
PR	27-SEP-1999; 99US-0156147.

THIS PAGE BLANK (USPTO)


```
Db      1 GGGGTCAACGTTGAGGGGG 20

RESULT 14
AX063578
LOCUS      AX063578                20 bp      DNA      linear      PAT 24-JAN-2001
DEFINITION Sequence 4 from Patent WO0100231.
ACCESSION  AX063578
VERSION     AX063578.1  GI:12541302
KEYWORDS    .
SOURCE      synthetic construct.
ORGANISM    synthetic construct.
REFERENCE   1 (bases 1 to 20)
AUTHORS     Cohen,J., Garcon,N. and Voss,G.
TITLE       Vaccines
JOURNAL     Patent: WO 0100231-A 4 04-JAN-2001;
            SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES    Location/Qualifiers
             1..20
             /organism="synthetic construct"
             /db_xref="taxon:32630"
             /note="oligonucleotide-nucleotides 1,2,15,16,17,18 and 19
             have linkages with a thioate modification"
BASE COUNT  3 a      2 c      12 g      3 t
ORIGIN

Query Match      84.0%; Score 16.8; DB 6; Length 20;
Best Local Similarity 90.0%; Pred. No. 3.3e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ggggtcacccgtgaggggg 20
        ||||| |..|||||
Db      1 GGGGTCAACGTTGAGGGGG 20

RESULT 15
AX088932
LOCUS      AX088932                20 bp      DNA      linear      PAT 17-MAR-2001
DEFINITION Sequence 4 from Patent WO0100232.
ACCESSION  AX088932
VERSION     AX088932.1  GI:13397690
KEYWORDS    .
SOURCE      synthetic construct.
ORGANISM    synthetic construct.
REFERENCE   1 (bases 1 to 20)
AUTHORS     Garcon,N. and Voss,G.
TITLE       Vaccine
JOURNAL     Patent: WO 0100232-A 4 04-JAN-2001;
            SmithKline Beecham Biologics SA (BE)
FEATURES    Location/Qualifiers
             1..20
             /organism="synthetic construct"
             /db_xref="taxon:32630"
             /note="CpG containing oligonucleotide"
BASE COUNT  3 a      2 c      12 g      3 t
ORIGIN

Query Match      84.0%; Score 16.8; DB 6; Length 20;
Best Local Similarity 90.0%; Pred. No. 3.3e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ggggtcacccgtgaggggg 20
        ||||| |..|||||
Db      1 GGGGTCAACGTTGAGGGGG 20
```

```
FT misc_feature /note="assembly_fragment"
FT 8294..12006
FT /note="assembly_fragment clone_end:SP6 vector_side:left"
FT 12107..16432
FT /note="assembly_fragment"
FT 16533..21956
FT /note="assembly_fragment"
FT 22057..28704
FT /note="assembly_fragment"
FT 28805..36083
FT /note="assembly_fragment"
FT 36184..47720
FT /note="assembly_fragment"
FT 47821..58804
FT /note="assembly_fragment"
FT 58905..73884
FT /note="assembly_fragment"
FT 73985..88989
FT /note="assembly_fragment"
FT 89090..104855
FT /note="assembly_fragment"
FT 104956..125506
FT /note="assembly_fragment"
FT 125607..160674
FT /note="assembly_fragment"
XX
SQ Sequence 160674 BP; 39338 A; 38937 C; 38154 G; 42844 T; 1401 other;

Query Match 85.0%; Score 17; DB 30; Length 160674;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggggtcacccggtgagg 17
|||||
Db 98517 GGGGTACCCGGTGAGG 98501

RESULT 11
AC105450/c
LOCUS AC105450 163782 bp DNA linear PRI 06-JAN-2002
DEFINITION Homo sapiens chromosome 2 clone RP11-112B11, complete sequence.
ACCESSION AC105450 AC060811
VERSION AC105450.1 GI:18072214
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 163782)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Waterston,R.H.
JOURNAL The sequence of Homo sapiens clone
REFERENCE 2 (bases 1 to 163782)
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
TITLE Submitted (06-JAN-2002) Genome Sequencing Center, Washington
UNIVERSITY School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jan 6, 2002 this sequence version replaced gi:8783741.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: H_NH0112B11
Drafting center: WBER
-----
Location/Qualifiers
1..163782
/organism="Homo sapiens"

FEATURES
source
```

```
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-112B11"
BASE COUNT 44805 a 36549 c 42489 g 39939 t
ORIGIN

Query Match 85.0%; Score 17; DB 9; Length 163782;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggggtcacccggtgagg 17
|||||
Db 46258 GGGGTACCCGGTGAGG 46242

RESULT 12
AR140453
LOCUS AR140453 20 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 12 from patent US 6207646.
ACCESSION AR140453
VERSION AR140453.1 GI:14482949
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg,A.M., Kline,J., Klinman,D. and Steinberg,A.D.
TITLE Immunostimulatory nucleic acid molecules
JOURNAL Patent: US 6207646-A 12 27-MAR-2001;
FEATURES Location/Qualifiers
1..20
source /organism="unknown"
BASE COUNT 3 a 2 c 12 g 3 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 20;
Best Local Similarity 90.0%; Pred. No. 3.3e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ggggtcacccggtgagg 20
|||||
Db 1 GGGGTACCGTTGAGGGGG 20

RESULT 13
AR154761
LOCUS AR154761 20 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 90 from patent US 6239116.
ACCESSION AR154761
VERSION AR154761.1 GI:15122814
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg,A.M. and Kline,J.N.
TITLE Immunostimulatory nucleic acid molecules
JOURNAL Patent: US 6239116-A 90 29-MAY-2001;
FEATURES Location/Qualifiers
1..20
source /organism="unknown"
BASE COUNT 3 a 2 c 12 g 3 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 20;
Best Local Similarity 90.0%; Pred. No. 3.3e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ggggtcacccggtgagg 20
|||||
Db 1 GGGGTACCGTTGAGGGGG 20
```

```

/ gene="TPO"
/ number=13
1508 .>2512
/ gene="TPO"
/ note="thyroid peroxidase intron M"
BASE COUNT      532 a   671 c   744 g   565 t
ORIGIN

Query Match      85.0%; Score 17; DB 9; Length 2512;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcaccgtgaggg 17
    |||
Db 1739 GGGTCACCGTGAGGG 1755

RESULT 10
AC060811/c
ID AC060811 standard; DNA; HTG; 160674 BP.
XX AC060811;
AC AC060811;
XX AC060811.3
SV AC060811.3
XX 24-APR-2000 (Rel. 63, Created)
DT 06-JUL-2000 (Rel. 64, Last updated, Version 3)
XX Homo sapiens chromosome 2 clone RP11-112B11 map 2, WORKING DRAFT SEQUENCE,
DE 15 unordered pieces.
XX HTG; HTGS_DRAFT; HTGS_PHASE1.
XX Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
XX [1]
XX Bliren B., Linton L., Nusbaum C., Lander E.;
RA "Homo sapiens chromosome 2, clone RP11-112B11";
RL Unpublished.
XX [2]
RA Bliren B., Linton L., Nusbaum C., Lander E., Abraham H., Allen N.,
RA Anderson S., Baldwin J., Barna N., Bastien V., Beda F., Boguslavskiy L.,
RA Boukhgalter B., Brown A., Burkett G., Campopiano A., Castie A., Choepel Y.,
RA Colangelo M., Collins S., Collamore A., Cooke P., DeArelano K., Dewar K.,
RA Diaz J.S., Dodge S., Domino M., Doyle M., Ferreira P., FitzHugh W.,
RA Gage D., Galagan J., Gardyna S., Ginde S., Goyette M., Graham L.,
RA Grand-Pierre N., Grant G., Hagos B., Heaford A., Horton L., Howland J.C.,
RA Iliev I., Johnson R., Jones C., Kann L., Karatas A., Klein J., LaRocque K.,
RA Lamazares R., Landers T., Lehoczy J., Levine R., Lieu C., Liu G.,
RA Locke K., MacDonald P., Marquis N., McCarthy M., McEwan P., McGurk A.,
RA McKernan K., McPheeters R., Meldrim J., Meneus L., Mihova T., Miranda C.,
RA Mienna V., Morrow J., Murphy T., Naylor J., Norman C.H., O'Connor T.,
RA O'Donnell P., O'Neill D., Oliver T.M., Oliver J., Peterson K., Pierre N.,
RA Pisan C., Pollara V., Raymond C., Riley R., Rogov P., Rothman D., Roy A.,
RA Santos R., Schauer S., Severy P., Spencer B., Stange-Thomann N.,
RA Stojanovic N., Subramanian A., Talamas J., Tesfaye S., Theodore J.,
RA Tirrell A., Travers M., Triggillo J., Vassiliev H., Viel R., Vo A.,
RA Wilson B., Wu X., Wyman D., Ye W.J., Young G., Zainoun J., Zimmer A.,
RA Zody M.;
RL Submitted (20-APR-2000) to the EMBL/GenBank/DBJ databases.
RL Whitehead Institute/MIT Center for Genome Research, 320 Charles Street,
RL Cambridge, MA 02141, USA
XX
CC On Jun 28, 2000 this sequence version replaced gi:8516129.
CC All repeats were identified using RepeatMasker:
CC Smit, A.F.A. & Green, P. (1996-1997)

```

```

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9800
Center clone name: 112_B.11
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 152611 bases at least Q40
Consensus quality: 156805 bases at least Q30
Consensus quality: 158332 bases at least Q20
Insert size: 168000; agarose-fp
Insert size: 159274; sum-of-contigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 4.7 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1886: contig of 1886 bp in length
* 1887 1986: gap of 100 bp
* 1987 4579: contig of 2593 bp in length
* 4580 4679: gap of 100 bp
* 4680 8193: contig of 3514 bp in length
* 8194 8293: gap of 100 bp
* 8294 12006: contig of 3713 bp in length
* 12007 12106: gap of 100 bp
* 12107 16432: contig of 4326 bp in length
* 16433 16532: gap of 100 bp
* 16533 21956: contig of 5424 bp in length
* 21957 22056: gap of 100 bp
* 22057 28704: contig of 6648 bp in length
* 28705 28804: gap of 100 bp
* 28805 36083: contig of 7279 bp in length
* 36084 36183: gap of 100 bp
* 36184 47720: contig of 11537 bp in length
* 47721 47820: gap of 100 bp
* 47821 58804: contig of 10984 bp in length
* 58805 58904: gap of 100 bp
* 58905 73884: contig of 14980 bp in length
* 73885 73984: gap of 100 bp
* 73985 88989: contig of 15005 bp in length
* 88990 89089: gap of 100 bp
* 89090 104855: contig of 15766 bp in length
* 104856 104955: gap of 100 bp
* 104956 125506: contig of 20551 bp in length
* 125507 125606: gap of 100 bp
* 125607 160674: contig of 35068 bp in length.
Key Location/Qualifiers
source 1. 160674
/db_chromosome="2"
/db_xref="taxon:9606"
/organism="Homo sapiens"
/map="2"
/clone="RP11-112B11"
/misc_feature 1. 1886
/note="assembly_fragment"
/misc_feature 1987. 4579
/note="assembly_fragment"
/misc_feature 4680. 8193

```

* 44300 47244: contig of 2945 bp in length
* 47245 47344: gap of unknown length
* 47345 51257: contig of 3913 bp in length
* 51258 51357: gap of unknown length
* 51358 54306: contig of 2949 bp in length
* 54307 54406: gap of unknown length
* 54407 58358: contig of 3952 bp in length
* 58359 58458: gap of unknown length
* 58459 61153: contig of 2695 bp in length
* 61154 61253: gap of unknown length
* 61254 64401: contig of 3148 bp in length
* 64402 64501: gap of unknown length
* 64502 67172: contig of 2671 bp in length
* 67173 67272: gap of unknown length
* 67273 70460: contig of 3188 bp in length
* 70461 70560: gap of unknown length
* 70561 72838: contig of 2278 bp in length
* 72839 72938: gap of unknown length
* 72939 75998: contig of 3060 bp in length
* 75999 76098: gap of unknown length
* 76099 79215: contig of 3117 bp in length
* 79216 82786: gap of unknown length
* 82787 82886: gap of unknown length
* 82887 86192: contig of 3306 bp in length
* 86193 86292: gap of unknown length
* 86293 89284: contig of 2992 bp in length
* 89285 89384: gap of unknown length
* 89385 91221: contig of 1837 bp in length
* 91222 91321: gap of unknown length
* 91322 94646: contig of 3325 bp in length
* 94647 94746: gap of unknown length
* 94747 97142: contig of 2396 bp in length
* 97143 97242: gap of unknown length
* 97243 99482: contig of 2240 bp in length
* 99483 99582: gap of unknown length
* 99583 102404: contig of 2822 bp in length
* 102405 102504: gap of unknown length
* 102505 104576: contig of 2072 bp in length
* 104577 104677: gap of unknown length
* 104677 107010: contig of 2334 bp in length
* 107011 107110: gap of unknown length
* 107111 108450: contig of 1340 bp in length
* 108451 108550: gap of unknown length
* 108551 110071: contig of 1521 bp in length
* 110072 110171: gap of unknown length
* 110172 112811: contig of 2640 bp in length
* 112812 112911: gap of unknown length
* 112912 115057: contig of 2146 bp in length
* 115058 115157: gap of unknown length
* 115158 117279: contig of 2122 bp in length
* 117280 117379: gap of unknown length
* 117380 119789: contig of 2410 bp in length
* 119790 119889: gap of unknown length
* 119890 122618: contig of 2729 bp in length
* 122619 122718: gap of unknown length
* 122719 123816: contig of 1098 bp in length
* 123817 123916: gap of unknown length
* 123917 126264: contig of 2348 bp in length
* 126265 126364: gap of unknown length
* 126365 127573: contig of 1209 bp in length
* 127574 127673: gap of unknown length
* 127674 129962: contig of 2289 bp in length
* 129963 130062: gap of unknown length
* 130063 132345: contig of 2283 bp in length
* 132346 132445: gap of unknown length
* 132446 133926: contig of 1481 bp in length
* 133927 134026: gap of unknown length
* 134027 136019: contig of 1993 bp in length
* 136020 136119: gap of unknown length
* 136120 137761: contig of 1642 bp in length
* 137762 137861: gap of unknown length
* 137862 140780: contig of 2919 bp in length

* 140781 140880: gap of unknown length
* 140881 142445: contig of 1565 bp in length
* 142446 142545: gap of unknown length
* 142546 144090: contig of 1545 bp in length
* 144091 144190: gap of unknown length
* 144191 145810: contig of 1620 bp in length
* 145811 145910: gap of unknown length
* 145911 147576: contig of 1666 bp in length
* 147577 147676: gap of unknown length
* 147677 149325: contig of 1649 bp in length
* 149326 149425: gap of unknown length
* 149426 150790: contig of 1365 bp in length
* 150791 150890: gap of unknown length
* 150891 152099: contig of 1209 bp in length
* 152100 152199: gap of unknown length
* 152200 154524: contig of 2325 bp in length
* 154525 154624: gap of unknown length
* 154625 156770: contig of 2146 bp in length
* 156771 156870: gap of unknown length
* 156871 158375: contig of 1505 bp in length
* 158376 158475: gap of unknown length
* 158476 159584: contig of 1109 bp in length
* 159585 159684: gap of unknown length

Query Match 87.0%; Score 17.4; DB 2; Length 180856;
Best Local Similarity 94.7%; Pred. No. 4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 99gtcacccggtgagggggg 20
||||| |||||||

Db 37410 GGGTCACTGGTGAGGGGG 37428

RESULT 9

HUMTPO11 Human thyroid peroxidase (TPO) gene, exons 12-13.
LOCUS HUMTPO11 2512 bp DNA linear
DEFINITION Human thyroid peroxidase (TPO) gene, exons 12-13.
ACCESSION M25711.1
VERSION M25711.1 GI:339859
KEYWORDS thyroid peroxidase.
SEGMENT 11 of 15
SOURCE Human lymphocyte DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2512)
AUTHORS Kimura,S., Hong,Y.S., Kotani,T., Ohtaki,S. and Kikkawa,F.
TITLE Structure of the human thyroid peroxidase gene: comparison and relationship to the human myeloperoxidase gene
JOURNAL Biochemistry 28 (10), 4481-4489 (1989)
MEDLINE 89352509
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted by S.Kimura,23-JUN-1989.
FEATURES
source Location/Qualifiers
1..2512
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2pter-p24"
<1..731
/gene="TPO"
/note="thyroid peroxidase Intron K"
prim_transcript <1..>2512
/gene="TPO"
/note="thyroid peroxidase mRNA and introns"
732..940
/gene="TPO"
/note="thyroid peroxidase, (EC 1.11.1.7); G00-120-446"
/number=12
941..1336
/gene="TPO"
/note="thyroid peroxidase Intron L"
1337..1507
exon

```

* 76781 78160: contig of 1380 bp in length
* 78161 78260: gap of unknown length
* 78261 79271: contig of 1011 bp in length
* 79272 79371: gap of unknown length
* 79372 80471 80570: contig of 1099 bp in length
* 80471 80570: gap of unknown length
* 80571 81748: contig of 1178 bp in length
* 81749 81848: gap of unknown length
* 81849 83322: contig of 1474 bp in length
* 83323 83422: gap of unknown length
* 83423 84851: contig of 1429 bp in length
* 84852 84951: gap of unknown length
* 84952 86314: contig of 1363 bp in length
* 86315 86414: gap of unknown length
* 86415 87465: contig of 1050 bp in length
* 87466 87564: gap of unknown length
* 87565 88595: contig of 1031 bp in length
* 88596 88695: gap of unknown length
* 88696 89771: contig of 1076 bp in length
* 89772 89871: gap of unknown length
* 89872 91004: contig of 1133 bp in length
* 91005 91105: gap of unknown length
* 91106 92136: contig of 1032 bp in length
* 92137 92236: gap of unknown length
* 92237 93294: contig of 1058 bp in length
* 93295 93394: gap of unknown length
* 93395 94685: contig of 1291 bp in length
* 94686 94785: gap of unknown length
* 94786 96171: contig of 1386 bp in length
* 96172 96271: gap of unknown length

Query Match
Best Local Similarity 94.7%; Pred. No. 4.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggtaccggtagggggg 19
|||||||
Db 25302 GGGGTACCCTGAGGGG 25320

RESULT 8
AC096062
LOCUS
DEFINITION
Rattus norvegicus chromosome SA clone CH230-59L13, *** SEQUENCING
IN PROGRESS ***, 71 unordered pieces.
AC096062
VERSION
AC096062.3 GI:17943732
KEYWORDS
HTG: HTGS_PHASE1.
SOURCE
Norway rat.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 180856)
Muzny,D.M., Adams,C.C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Aisbrooks,S.L., Anarstange,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Boucek,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhaq,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyte,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,
Hollins,B., Homsy,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseghe,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,A., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,M., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogih,M., Okwuonu,G., Oranuye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 180856)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:16901715.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hpsc-help@bcm.tmc.edu
----- Project Information
Center project name: GEIM
Center clone name: CH230-59L13
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 163243 bases at least Q40
Consensus quality: 168450 bases at least Q30
Consensus quality: 172632 bases at least Q20
Estimated insert size: 161344; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 71 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 7525: contig of 7525 bp in length
* 7526 7625: gap of unknown length
* 7626 17453: contig of 9828 bp in length
* 17454 17553: gap of unknown length
* 17554 24023: contig of 6470 bp in length
* 24024 24123: gap of unknown length
* 24124 28990: contig of 4867 bp in length
* 28991 29090: gap of unknown length
* 29091 31249: contig of 2159 bp in length
* 31250 31349: gap of unknown length
* 31350 35615: contig of 4266 bp in length
* 35616 35715: gap of unknown length
* 35716 40568: contig of 4853 bp in length
* 40569 40668: gap of unknown length
* 40669 44199: contig of 3531 bp in length
* 44200 44299: gap of unknown length

```

Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
 Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S.,
 Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
 Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C.,
 Hollins, B., Homs, F., Howard, S., Huber, J., Huly, S., Hume, J.,
 Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
 Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
 Kovach, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C.,
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
 Louised, H., Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
 Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
 Martinez, E., Massey, E., Mawhinney, E., McLeod, M. P., Meador, M.,
 Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
 Morgan, A., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
 Nguyen, M., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkwo, S.,
 Ogulu, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
 Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L.,
 Quiles, M., Ruiz, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,
 Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshtari, N.,
 Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H.,
 Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N.,
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R.,
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
 Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,
 Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 100587)
 Worley, K. C.
 Direct Submission
 Submitted (05-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Dec 20, 2001 this sequence version replaced gi:17066816.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GGXK
 Center clone name: CH230-88F22
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329First call to
 findPhraplist
 Consensus quality: 68083 bases at least Q40
 Consensus quality: 75814 bases at least Q30
 Consensus quality: 82939 bases at least Q20
 Estimated insert size: 61693; sum-of-contigs estimation
 Quality coverage: 0x in Q20 bases; agarose-fp estimation
 Quality coverage: 0.6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 59 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 * 1 2812: contig of 2812 bp in length
 * 2813 2912: gap of unknown length
 * 2913 5750: contig of 2838 bp in length
 * 5751 5850: gap of unknown length
 * 5851 8558: contig of 2708 bp in length
 * 8559 8659: gap of unknown length
 * 8659 11205: contig of 2547 bp in length
 * 11205 11305: gap of unknown length
 * 11306 13239: contig of 1934 bp in length

13240 13339: gap of unknown length
 13340 15119: contig of 1980 bp in length
 15119 15419: gap of unknown length
 15420 17657: contig of 2238 bp in length
 17658 17757: gap of unknown length
 17758 21225: contig of 3468 bp in length
 21226 21325: gap of unknown length
 21326 23259: contig of 1934 bp in length
 23260 23359: gap of unknown length
 23360 24598: contig of 1339 bp in length
 24599 26718: gap of unknown length
 26719 26818: contig of 1920 bp in length
 26799 28361: gap of unknown length
 28362 28461: gap of unknown length
 28462 29807: contig of 1346 bp in length
 29808 32794: contig of 2887 bp in length
 32795 32894: gap of unknown length
 32895 34921: contig of 2027 bp in length
 34922 35021: gap of unknown length
 35022 36223: contig of 1202 bp in length
 36224 38580: contig of 2257 bp in length
 38581 38680: gap of unknown length
 38681 40515: contig of 1835 bp in length
 40516 41760: gap of unknown length
 41761 41860: gap of unknown length
 41861 44103: contig of 2243 bp in length
 44104 44203: gap of unknown length
 44204 45855: contig of 1652 bp in length
 45856 45955: gap of unknown length
 45956 47795: contig of 1840 bp in length
 47796 47896: gap of unknown length
 47897 49924: contig of 2029 bp in length
 49925 51408: contig of 1384 bp in length
 51409 51508: gap of unknown length
 51509 53142: contig of 1634 bp in length
 53143 53442: gap of unknown length
 53443 54483: contig of 1241 bp in length
 54484 54583: gap of unknown length
 54584 55809: contig of 1226 bp in length
 55810 55909: gap of unknown length
 55910 57462: contig of 1553 bp in length
 57463 59265: gap of unknown length
 59266 59365: gap of unknown length
 59366 60582: contig of 1217 bp in length
 60583 61919: contig of 1237 bp in length
 61920 62019: gap of unknown length
 62020 63627: contig of 1608 bp in length
 63628 63727: gap of unknown length
 63729 65379: contig of 1652 bp in length
 65380 65479: gap of unknown length
 65480 67211: contig of 1732 bp in length
 67212 68449: contig of 1138 bp in length
 68450 68549: gap of unknown length
 68550 70104: contig of 1555 bp in length
 70105 70204: gap of unknown length
 70205 71455: contig of 1251 bp in length
 71456 71556: gap of unknown length
 71557 72798: contig of 1143 bp in length
 72799 73993: contig of 1195 bp in length
 73994 74093: gap of unknown length
 74094 75516: contig of 1422 bp in length
 75517 76680: contig of 1065 bp in length
 76681 76780: gap of unknown length

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT


```

Query Match      100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
    |||
Db 1 ggggtcacccggtgagggggg 20

RESULT 2
LOCUS AX104851
DEFINITION Sequence 1043 from Patent WO0122972.
ACCESSION AX104851
VERSION AX104851.1 GI:13921048
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial construct
REFERENCE Krieg,A.M., Schetter,C. and Vollmer,J.C.
AUTHORS Immunostimulatory nucleic acids
TITLE Patent: WO 0122972-A 1043 05-APR-2001;
JOURNAL UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES
    source
        1. .20
            /organism="synthetic construct"
            /db_xref="taxon:32630"
BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match      100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
    |||
Db 1 ggggtcacccggtgagggggg 20

RESULT 2
LOCUS AX104851
DEFINITION Sequence 1043 from Patent WO0122972.
ACCESSION AX104851
VERSION AX104851.1 GI:13921048
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial construct
REFERENCE Krieg,A.M., Schetter,C. and Vollmer,J.C.
AUTHORS Immunostimulatory nucleic acids
TITLE Patent: WO 0122972-A 1043 05-APR-2001;
JOURNAL UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES
    source
        1. .20
            /organism="synthetic construct"
            /db_xref="taxon:32630"
BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match      100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
    |||
Db 1 ggggtcacccggtgagggggg 20

RESULT 3
LOCUS AX105126
DEFINITION Sequence 24 from Patent WO0122990.
ACCESSION AX105126
VERSION AX105126.1 GI:13921276
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial construct
REFERENCE Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
AUTHORS Methods related to immunostimulatory nucleic acid-induced
TITLE interferon
JOURNAL Patent: WO 0122990-A 24 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
    source
        1. .20
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="Synthetic Oligonucleotide"
        misc_feature 1. .2
            /note="Backbone has phosphorothioate linkages."
        misc_feature 3. .14
            /note="Backbone has phosphodiester linkages."
        misc_feature 15. .19
            /note="Backbone has phosphorothioate linkages."
        misc_feature 20
            /note="Backbone has phosphodiester linkages."

```

```

BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match      100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggtcacccggtgagggggg 20
    |||
Db 1 ggggtcacccggtgagggggg 20

RESULT 4
LOCUS AP004675/c
DEFINITION Oryza sativa chromosome 8 clone P0709D11, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION AP004675
VERSION AP004675.1 GI:18307753
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:P0709D11.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 146884)
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Direct Submission
JOURNAL Submitted (23-JAN-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
    source
        1. .146884
            /organism="Oryza sativa"
            /cultivar="Nipponbare"
            /db_xref="taxon:4530"
            /chromosome="8"
            /clone="P0709D11"
BASE COUNT 41459 a 31135 c 31177 g 43057 t 56 others
ORIGIN

Query Match      90.0%; Score 18; DB 2; Length 146884;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ggtcacccggtgagggggg 20
    |||
Db 15214 GGTACCGGTGAGGGGGG 15197

RESULT 5
LOCUS AB037276
DEFINITION Cynomolgus Epstein-Barr Virus SI-IIA gene for EBNA-1, complete cds.
ACCESSION AB037276
VERSION AB037276.1 GI:9453858
KEYWORDS EBNA-1.
SOURCE Cynomolgus Epstein-Barr Virus SI-IIA cell_line:SI-IIA DNA.
ORGANISM Cynomolgus Epstein-Barr Virus SI-IIA

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:58:07 ; Search time 2778.35 seconds
(without alignments)
150.640 Million cell updates/sec

Title: US-09-672-126-24
Perfect score: 20
Sequence: 1 ggggtcacccgtgagggggg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

1	20	100.0	20	6	AX104787	Sequence
2	20	100.0	20	6	AX104851	Sequence
3	20	100.0	20	6	AX105126	Sequence
c 4	18	90.0	145884	2	AF004675	Oryza sat
5	17.4	87.0	2259	14	AB037276	Cynomolgu
c 6	17.4	87.0	76072	8	NCB1D1	AL355927 Neurospor
7	17.4	87.0	100587	2	AC096978	Rattus no
8	17.4	87.0	180856	2	AC096062	AC096062 Rattus no
9	17	85.0	2512	9	HUMTPO11	M25711 Human thyro
c 10	17	85.0	166674	30	AC060811	AC060811 Homo sapi
c 11	17	85.0	163782	9	AC105450	AC105450 Homo sapi
12	16.8	84.0	20	6	AR140453	AR140453 Sequence
13	16.8	84.0	20	6	AR154761	AR154761 Sequence
14	16.8	84.0	20	6	AX063578	AX063578 Sequence
15	16.8	84.0	20	6	AX088932	AX088932 Sequence
16	16.8	84.0	20	6	AX104327	AX104327 Sequence
17	16.8	84.0	20	6	AX104575	AX104575 Sequence
18	16.8	84.0	20	6	AX104776	AX104776 Sequence
19	16.8	84.0	20	6	AX104777	AX104777 Sequence
20	16.8	84.0	20	6	AX105103	AX105103 Sequence
21	16.8	84.0	20	6	AX105108	AX105108 Sequence
22	16.8	84.0	20	6	AX105236	AX105236 Sequence
23	16.8	84.0	20	6	AX135634	AX135634 Sequence
24	16.8	84.0	20	6	AX194489	AX194489 Sequence
25	16.8	84.0	20	6	AX355408	AX355408 Sequence
26	16.8	84.0	20	6	AX355409	AX355409 Sequence
27	16.8	84.0	20	6	BD009060	BD009060 Immunosti
28	16.8	84.0	21	6	AX104755	AX104755 Sequence
29	16.8	84.0	21	6	AX104811	AX104811 Sequence
30	16.8	84.0	21	6	AX104812	AX104812 Sequence
31	16.8	84.0	21	6	AX105119	AX105119 Sequence
32	16.8	84.0	21	6	AX105257	AX105257 Sequence
33	16.8	84.0	24	6	AX104326	AX104326 Sequence
c 34	16.8	84.0	4080	4	OSAINFGFII3	U00664 Ovis aries
35	16.8	84.0	5317	9	HSMB02401	AL137634 Homo sapi
36	16.8	84.0	8094	10	MUSAP	D16195 Mouse gene
c 37	16.8	84.0	10575	1	AE000917	AE000917 Methanoba
38	16.8	84.0	12637	1	AF323753	AF323753 Streptomy
c 39	16.8	84.0	19391	10	MMDSMINP	Z18892 Mus musculu
c 40	16.8	84.0	27541	6	AX211706	AX211706 Sequence
c 41	16.8	84.0	39324	9	HSL27H9	Z49237 Human DNA f
c 42	16.8	84.0	57351	2	AC095917	AC095917 Rattus no
c 43	16.8	84.0	64122	2	AC097606	AC097606 Rattus no
44	16.8	84.0	71553	2	AC096997	AC096997 Takifugu
c 45	16.8	84.0	74138	2	AC021272	AC021272 Homo sapi

ALIGNMENTS

RESULT 1	AX104787	Sequence 979 from Patent WO0122972.	20 bp	DNA	linear	PAT 30-APR-2001
LOCUS	AX104787	Sequence 979 from Patent WO0122972.				
DEFINITION	AX104787					
ACCESSION	AX104787.1	GI:13920984				
VERSION	AX104787.1	GI:13920984				
KEYWORDS						
SOURCE		synthetic construct.				
ORGANISM		artificial sequence.				
REFERENCE		1 (bases 1 to 20)				
AUTHORS		Krieg, A.M., Schetter, C. and Vollmer, J.C.				
TITLE		Immunostimulatory nucleic acids				
JOURNAL		Patent: WO 0122972-A 979 05-APR-2001;				
		UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical				
FEATURES		GmbH (DE)				
Source		Location/Qualifiers				
		1..20				
		/organism="synthetic construct"				
		/db_xref="taxon:32630"				
BASE COUNT		2 a	3 c	13 g	2 t	
ORIGIN						

Search completed: August 10, 2002, 03:06:09
Job time: 16035 sec

EARLIER APPLICATION NUMBER: 60/094,783
; EARLIER FILING DATE: JULY 31, 1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1694
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-362-473-3

Query Match 71.0%; Score 14.2; DB 4; Length 1694;
Best Local Similarity 84.2%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ggggacgacgttggtggggg 20
||||| ||| ||| ||| ||| |||
DB 477 GGGGAGGATGTTGGGGGG 459

RESULT 15
US-09-155-036-3
; Sequence 3, Application US/09155036
; Patent No. 6265201
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF MINNESOTA
; TITLE OF INVENTION: DNA MOLECULES AND PROTEIN DISPLAYING
; TITLE OF INVENTION: IMPROVED TRIAZINE COMPOUND DEGRADING ABILITY
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
; STREET: 119 No. 6265201th Fourth Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155,036
; FILING DATE: 16-JAN-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/035,404
; FILING DATE: 17-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCORMACK, MYRA M.
; REGISTRATION NUMBER: 36,602
; REFERENCE/DOCKET NUMBER: 110.00400201
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1225
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1698 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-155-036-3

Query Match 71.0%; Score 14.2; DB 4; Length 1808;
Best Local Similarity 84.2%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ggggacgacgttggtggggg 20
||||| ||| ||| ||| ||| |||
DB 1287 GGGGACGATGTTGGGGG 1305

misc_feature /note="Pfam match to entry PF01574 IMPDH_N, IMP dehydrogenase / GMP reductase N terminus, score 175.90, E-value 6.8e-49", 1333..2411

misc_feature /gene="SCD63.02" /note="Previously sequenced DNA fragment EMBL:AJ010601 Streptomyces coelicolor A3(2) DNA for whiD and whiK loci" 1349..1510

misc_feature /gene="SCD63.02" /note="Pfam match to entry PF00571 CBS, CBS domain, score 46.80, E-value 5e-10" 1532..1693

misc_feature /gene="SCD63.02" /note="Pfam match to entry PF00571 CBS, CBS domain, score 40.00, E-value 5.4e-08" 1781..2464

misc_feature /gene="SCD63.02" /note="Pfam match to entry PF00478 IMPDH_C, IMP dehydrogenase / GMP reductase C terminus, score 338.30, E-value 8.7e-98" 1955..1993

misc_feature /gene="SCD63.02" /note="PS00487 IMP dehydrogenase / GMP reductase signature" 2425..2940

misc_feature /note="Previously sequenced DNA fragment EMBL:AJ010601 Streptomyces coelicolor A3(2) DNA for whiD and whiK loci" 2701..3825

gene /gene="SCD63.03" /note="SCD63.03, possible inosine-5'-monophosphate dehydrogenase, len: 374 aa; N-terminal region identical to previously sequenced TR:086845 (EMBL:AJ010601) Streptomyces coelicolor hypothetical; 38.9 kD protein (fragment), 64 aa and whole CDS similar to SW:IMDH_ECOLI (EMBL:X02209) Escherichia coli inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205) GubB, 488 aa; fasta scores: opt: 275 z-score: 311.9 E(): 6.1e-10; 29.0% identity in 376 aa overlap. Contains Pfam matches to entries PF01574 IMPDH_N, IMP dehydrogenase / GMP reductase N terminus PF00478 IMPDH_C, IMP dehydrogenase / GMP reductase C terminus" /codon_start=1 /transl_table=11 /product="putative inosine-5'-monophosphate dehydrogenase" /protein_id="G1:7320890" /db_xref="GI:7320890" /translation="MTEIIGKGRGRRAVAFDIAVPSRRTRDPEKSIAMQIDAY RELPLIAPMDSVSPATAIRIGELGGLVLEGLMTLRHEDPQLDELIALDIDN ATRRLQEIYAAPIKELLIGORIKEDVSGVTPAAQSPQAFSAKAVDAGDIVI RGTVAEHSVSGSHEPLNKKOYELDVPYVGGCATYAAHLMTGAAGVYVKG GAHTTRNVLGIVPMATYADVAAARRRDYSGSRRYVHVDGVSNGSGLPKALA CGADSVWMSPLARATDAPGRGNHGMENVEELPGKRVVDLGTGTIELTGPRSN PDGSMNFALRAAMATGYSELKEFRVEVYVADSQHR" 2725..2934

misc_feature /gene="SCD63.03" /note="Pfam match to entry PF01574 IMPDH_N, IMP dehydrogenase / GMP reductase N terminus, score 1.10, E-value 0.00079" 3145..3792

misc_feature /gene="SCD63.03" /note="Pfam match to entry PF00478 IMPDH_C, IMP dehydrogenase / GMP reductase C terminus, score 44.60, E-value 1.1e-09" 3450..3566

misc_feature /gene="SCD63.03" /note="Previously sequenced DNA fragment EMBL:AJ010601 Streptomyces coelicolor A3(2) DNA for whiD and whiK loci" 4157..5377

gene /gene="SCD63.04" /note="SCD63.04" 4157..5377

CDS /gene="SCD63.04" /gene="SCD63.04"

Query Match 76.3%; Score 20.6; DB 1; Length 31624; Best Local Similarity 85.2%; Pred. No. 2.2e+03; Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gggagtcagctgcagctcagggggg 27
DB 27563 GGCCTGCCCTGCACGCCGAGGAGGG 27589

RESULT 13
AC005929
LOCUS 41944 bp DNA linear INV 24-FEB-2000
DEFINITION Leishmania major chromosome 3 clone L6910 strain Friedlin, complete sequence.
ACCESSION AC005929
VERSION AC005929.5 GI:7025830
KEYWORDS HTG.
SOURCE Leishmania major.
ORGANISM Leishmania major
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

REFERENCE 1 (bases 1 to 41944)
AUTHORS Myler,P.J., Sisk,E., Hixson,G., Kiser,P., Rickel,E., Hassebrock,M., Cawthra,J., Marsolini,F., Sunkin,S. and Stuart,K.D.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-1998) Seattle Biomedical Research Institution, 4 Nickerson Street, Seattle, WA 98109-1651, USA

REFERENCE 2 (bases 1 to 41944)
AUTHORS Myler,P.J., Sisk,E., Hixson,G., Kiser,P., Rickel,E., Hassebrock,M., Cawthra,J., Marsolini,F., Sunkin,S. and Stuart,K.D.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-1999) Seattle Biomedical Research Institution, 4 Nickerson Street, Seattle, WA 98109-1651, USA

REFERENCE 3 (bases 1 to 41944)
AUTHORS Myler,P.J., Sisk,E., Hixson,G., Kiser,P., Rickel,E., Hassebrock,M., Cawthra,J., Marsolini,F., Sunkin,S. and Stuart,K.D.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2000) Seattle Biomedical Research Institution, 4 Nickerson Street, Seattle, WA 98109-1651, USA
COMMENT On Feb 24, 2000 this sequence version replaced gi:6425645.

FEATURES
source location/Qualifiers
1..41944
/organism="Leishmania major"
/strain="Friedlin"
/db_xref="taxon:5664"
/chromosome="3"
/clone="L6910"
1420..2418
/gene="GOR"
/note="L6910.13; predicted using Glimmer, Testcode and CodonUsage"
1420..2418
/gene="GOR"
/codon_start=1
/product="Quinone Oxidoreductase"
/protein_id="AF35929.1"
/db_xref="GI:7025831"
/translation="MEVIADAHGPEVLAVRPSSTPDATOLEGGQVLAHNNVAGVNF IDYVFLSGLYKKRPAMYIYGEAGAVVVGAGVPTMTGKRVAYRGGAGCGSYVAF TVAPASLREVPDGVTDAAARAAMCGLTRHTLYVDSYTCGSGESTYVYVRAAGGTLV VCOMARKGARVYDGCAGAEKATLARSVGRADVIDYVATPDAPLVRAADQGVDAV YDVGQARTFAGSLSVLRPGMYITTEGNASGAVPPVPLESRAGSVYLDRPTLGDMR TPEEVGARTSEVEFWAASROVOLTIGHEVRLHBAQALDLSRKTGKLLKCID" 2880..3416
/gene="L6910.1"
/note="predicted using Glimmer, Testcode and CodonUsage"
2880..3416
/gene="L6910.1"

gene /gene="L6910.1" /gene="L6910.1"

CDS


```

/translation="MSDALDAFHDHDAFTALGTAVSYALVLYMTAVLFGVPIAETFA
I."
complement(8300. .9406)
/sequence="
/sequence="prirv2"
/sequence="VNG05106"
complement(8300. .9406)
/sequence="prirv2"

```

Query Match	76.3%	Score 20.6;	DB 1;	Length 14379;
Best Local Similarity	85.2%	Pred. No. 2.5e+03;		
Matches 23; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY	1	ggggtcgcgtcgacgtcgaggggg	27
Db	7470	GGCGACGACGTGCATCGAGGAGGG	7444

FEATURE	LOCATION	COORDINATES	ORIENTATION	FEATURE	LOCATION	COORDINATES	ORIENTATION
LOCUS	SCD63	31624 bp	DNA	LOCUS	SCD63	31624 bp	DNA
DEFINITION	Streptomyces coelicolor cosmid b63.		linear	DEFINITION	Streptomyces coelicolor cosmid b63.		linear
ACCESSION	AL161755			ACCESSION	AL161755		
VERSION	AL161755.1			VERSION	AL161755.1		
KEYWORDS	aldehyde dehydrogenase; aldolase; cholesterol oxidase; ECF sigma			KEYWORDS	aldehyde dehydrogenase; aldolase; cholesterol oxidase; ECF sigma		

SOURCE ORGANISM	REFERENCE	TITLE	JOURNAL
Streptomyces coelicolor A3(2).	Redenbach, M., Kiese, H.M., Denaplatz, D., Eichner, A., Cullum, J., Kinashi, H., and Hopwood, D.A.	A set of ordered cosmids and a detailed genetic and physical map for the 8 mb Streptomyces coelicolor A3(2) chromosome	Mol. Microbiol. 21 (1), 77-96 (1996)
Streptomyces coelicolor A3(2).	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces 1 (bases 1 to 31624)		

COMMENTS

Notes: Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC and Beowulf Genomics. Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous.

program of `libd` et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov/jp/jun/cg1-bin/frameplot.pl>. **CAUTION:** We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (`atg`, `gtg`, `tig` or `atc`) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions.

Cosmid D63.

FEATURES
SOURCE

misc_feature

misc_feature

gene

CDS

misc_feature

gene

CDS

misc_feature

```

www.ncbi.nlm.nih.gov/.../Gene_30:157-66(1984) as implemented at
www.rnh.go.jp/...
bin/frameplot.pl. CAUTION: We may not have predicted the
initiation codon. Where possible we choose an upstream
gag, gtg, ttg or (att) which is preceded by an initiation
binding site sequence (optimally 5-13bp before the
start codon). If this cannot be identified we choose the most
likely initiation codon.
1. This sequence MAY NOT be the entire insert of the
cloned clone. It may be shorter because we only sequence
certain sections once, or longer, because we arrange for a
partial overlap between neighbouring submissions.
53.
Location/Qualifiers
1. 31624
/organism="Streptomyces coelicolor A3(2)"
/strain="A3(2)"
/db_xref="taxon:100226"
/clone="cosmid D63"
1. 98
/note="nominal overlap with S. coelicolor cosmid SC664"
1. 947
/note="Previously sequenced fragment EMBL:AJ010601
Streptomyces coelicolor A3(2) DNA for whid and whik loci"
328. 915
/gene="SCD63.01"
328. 915
/gene="SCD63.01"
/note="SCD63.01", ECF sigma factor, len: 195 aa; identical
to previously sequenced TR:086843 (EMBL:AJ010601)
Streptomyces coelicolor ECF sigma factor, 195 aa"
/codon_start=1
/transl_table=11
/product="ECF sigma factor"
/protein_id="CAB82008.1"
/db_xref="GI:7320888"
/translation="MPRDAAPPGVGSTVHRAVDEGDAATHDIAVHPILARYCR
TRSLRFGDARHVEDLAQVCYCAVLAALPRKDRGPEFAVPAIAHKKADQGRRA
MRHGSTRAVSDPEKPEPDSCTEPEERALLNSDAAMAKKLANIPENORELLRLIAV
GLTAETGTMGMSPGAVRAOHRALSRLLALAEQ"
1025. 1145
/note="Previously sequenced DNA fragment EMBL:AJ010601
Streptomyces coelicolor A3(2) DNA for whid and whik loci"
1055. 2560
/gene="SCD63.02"
1055. 2560
/gene="SCD63.02"
/note="SCD63.02", inosine 5' monophosphate dehydrogenase,
len: 501 aa; highly similar to previously sequenced
TR:086844 (EMBL:AJ010601) Streptomyces coelicolor inosine
5' monophosphate dehydrogenase, 523 aa; fasta scores: opt:
2626 z-score: 2896.3 E(): 0; 84.9% identity in 509 aa
overlap. Contains Pfam matches to entries PF01574 IMPDH_N,
CBP dehydrogenase / GMP reductase N terminus, 2x PF00571
CBP, CBS domain and PF00478 IMPDH_C, IMP dehydrogenase /
GMP reductase C terminus and match to Prosite entry
PS00487 IMP dehydrogenase / GMP reductase signature"
/codon_start=1
/transl_table=11
/product="inosine 5' monophosphate dehydrogenase"
/protein_id="CAB82009.1"
/db_xref="GI:7320889"
/transl_table=11
/organism="MpanDGVGEKFAFLGLTYDDVLLPGASAVLPANVADTSSRISR
NVRNIPILLASAMKQVESHMSAIAKQSGVAGVLLRLNLSIDQANQVDLYRSRGVAV
ANPTIHPDALTGADALCKKPIISCVPTDQVGLGLVNRNIDGVLKPPVLRREV
MTPEPLITGVGVGSGVSDPMLRHRKIKITLYDDDGITLKGLTYKQDVRAEYQPHAA
KDAKGRLLVGAAGASPEALDRQAQLAEQVDFVDTTSHGNSNALSMSKTISSVLA
IDVGGNVAIRDAQALIDAGVGIKVGVGPSICTTTRVAGIGVPPVATLSSVLA
RAACVPIIGCGGQIGSDGIDKALAAAGADVMGLSLACGESPGLFINCKQKSYR
SGSLGMSGSGGGRSGSKRYQVAADVDMLPEEGIEGVGPAGPLANYLHOLVGG
LRQRMGVGAATIEEMSKRFRITISALIKSHRHDIQMIVTEAPNVSRSK"
1082. 1345
/gene="SCD63.02"

```

```

CDS
79. .825
/ gene="pod"
/ note="pod"
/ codon_start=1
/ transl_table=11
/ product="3-isopropylmalate dehydratase"
/ protein_id="AAG19032.1"
/ db_xref="GI:10580104"
/ translation="MGSVTHTTAEDIEDPVLIGLPGVGLVGRILANDHITIDVDMSCV
GVSACAIPQVTVYTEDSHAAIPPRVRIYADESGKILLALOGDPIISRNGEPRGTCYTG
WIDHNVTPYIISGLPSDDHELDLPVTEIGSGALPERIAALDITGPPRGVITGGT
GALMTATQRLDAAALVVEDSPQFPDPAGARLLIDHAIPLVEVSPITDLDVDRAD
IRSQEQLEQMQHDEHSTQAKPLRMFG"
complement(836. .2074)
/ gene="yfmO1"
/ note="yfmO1"
complement(836. .2074)
/ gene="yfmO1"
/ note="yfmO1"
/ codon_start=1
/ transl_table=11
/ product="multidrug resistance protein homolog"
/ protein_id="AAG19033.1"
/ db_xref="GI:10580105"
/ translation="NVRRAHMYMLSGSLGPPSPHQMGFRDFHPAVLATATATETAF
LGIGVDPILPSIAKELGADHFMWFLTSYLTGTGLANVAGTLATIRGSKNTMVLG
IGIVAVAGGALVSSVESMAVLRGFMGLGNALFTTTALAIIIVGSAGNGRAITTYE
AALRGSIAGGPIIGLGLCAQSMRLPFGAGTSVLMILGFVPATTVTEPDTEGRGVRL
IGTRHTGVRINLIVGMLTYTGFTVLAIVPLVEGTTMQIGLVFFAMGLIVASA
VLAPRLNARFGTPEVASAALTAFAALCLMIVLWISVPLVAGLVTAAGLGLNANIS
TLAMGISTHSRPSAGVNSLRTGGAFAPIVAGYIGVNGALPPLGAAVITLGI
VLLARSAFTTATPPADVGDH"
2239. .3360
/ gene="aspB1"
/ note="VNG0502G"
2239. .3360
/ gene="aspB1"
/ note="aspB1"
/ codon_start=1
/ transl_table=11
/ product="aspartate aminotransferase"
/ protein_id="AAG19034.1"
/ db_xref="GI:10580106"
/ translation="MTEFSDRVAVQVSIIGIRAFVEAAGEDAINLGLGDEPTPDHAR
QAADVADIESGADGYSNRGTALVDAIYEKHAQDGVAPAGVATATAGGSPALLA
MEAHNDPGDEVLEPPRGVSYDALTRMAGNPGPLRPDDTLTAPETVDDHITDDAA
FYVNSPANTGAVQSPADMRARIRIDEDHDLICISDEYEHIVFESEHNSPFEFADTD
NVVVVACSKTYSMTGWRGLDGLAEMGLDVPTEGAFYAMPTEVDDGWIDEVDRGVVV
PGSAFGDHGAGTARISYATDMATLIRDAIDVMAATAAVQ"
3425. .4213
/ gene="VNG0503C"
3425. .4213
/ gene="VNG0503C"
/ note="conserved hypothetical protein"
/ codon_start=1
/ transl_table=11
/ product="Vng0503c"
/ protein_id="AAG19035.1"
/ db_xref="GI:10580107"
/ translation="MTDQADDDVVAQVADGGILLARALDAFRTGERSHWEMLFDH
LDLGPAAVLISGAGHAGFWANADRPDGMVDYTDAGVADAMDALADAPHEFT
FVVVDARQIPYADTFDVAHHLVLAHLGDADRRQALAEIARVLPPTGELHAATSGDD
ALTELMAVARHGDHVDPERPTLONCAQLRDFEALVVERRRDLAPLVNAPRELVAVY
LSLPGDFPAGALERAFAFRARYDGLTIVTGDGKVLIARHPDTA"
complement(4210. .5391)
/ gene="gdp5"
/ note="VNG0504G"
complement(4210. .5391)
/ gene="gdp5"
/ note="gdp5"
/ codon_start=1
/ transl_table=11

gene
502. .6602
/ gene="VNG0505C"
502. .6602
/ gene="VNG0505C"
/ note="conserved hypothetical protein"
/ codon_start=1
/ transl_table=11
/ product="Vng0505c"
/ protein_id="AAG19037.1"
/ db_xref="GI:10580109"
/ translation="MSAGETTPYTAIRIRGIALLAGNGEFVATFETALTPVPDESA
FQETAGLRRAEWTDRIETITLTCQGLITVCSGLGVAEPPALAVLTPVEEVFGA
GSIATLSVLAIVMVAHLVLAEOAPTYLIGRAKTAHYCAPLHWRTSPITIR
AGDIRIAKILISLVGEMTSRTEGAEDPDGSRGDLISGTGPDDEBEVY
GALDIRIPYGDVAIVDEVAAYSTDDLEPTNDRADSPPHRYPLVGESEMDFERGV
YLPEARDDLDLRDGLTALAEVVAHPLVFPADMSDAIDRFQAEQOLAFTVDADTG
GVVGWYITDYLATIGVEEDPTDEAAKSQS"
complement(6599. .6730)
/ gene="VNG0506H"
complement(6599. .6730)
/ gene="VNG0506H"
/ note="hypothetical protein"
/ codon_start=1
/ transl_table=11
/ product="Vng0506h"
/ protein_id="AAG19038.1"
/ db_xref="GI:10580110"
/ translation="MGEPPTPAAGYGVSLMDALSALVFGVYFLTVNAGSAARLD"
complement(6931. .7803)
/ gene="VNG0507C"
complement(6931. .7803)
/ gene="VNG0507C"
/ note="conserved hypothetical protein"
/ codon_start=1
/ transl_table=11
/ product="Vng0507c"
/ protein_id="AAG19039.1"
/ db_xref="GI:10580111"
/ translation="MPMRRLDCGVLPPTIDOLHIPTGTTVEKHDVADSVLIGAS
QURLGVGHNVVAGSERVSDGIEAAGDQRLDMWCVTGTVLVEGSAVYGERHTIER
IYVGDLDIDVDYDIEGFEASGWIYTRDMPRLTFYGNITLHLILGAEADQELVE
ELAADDRCGRPLTIPQASVSDDAWRVSTPATIGDCCRLGNIRATITDVGRCGNVVS
LRAQNDVTVGAGTKIHGDVYTRNGSVHVGODAVVLDVSGHDVTIDEADVDSVIRAR
GEMRVGSVADRDE"
7872. .8069
/ gene="VNG0508H"
7872. .8069
/ gene="VNG0508H"
/ note="hypothetical protein"
/ codon_start=1
/ transl_table=11
/ product="Vng0508h"
/ protein_id="AAG19040.1"
/ db_xref="GI:10580112"
/ translation="MTSLSVADGDVVYEGTEFLERLEKPLIEDATQSDHYVTDHDL
KLVEPNPTLSGGERRIGITLD"
8119. .8256
/ gene="VNG0509H"
8119. .8256
/ gene="VNG0509H"
/ note="hypothetical protein"
/ codon_start=1

gene
502. .6602
/ gene="VNG0505C"
502. .6602
/ gene="VNG0505C"
/ note="conserved hypothetical protein"
/ codon_start=1
/ transl_table=11
/ product="Vng0505c"
/ protein_id="AAG19037.1"
/ db_xref="GI:10580109"
/ translation="MSAGETTPYTAIRIRGIALLAGNGEFVATFETALTPVPDESA
FQETAGLRRAEWTDRIETITLTCQGLITVCSGLGVAEPPALAVLTPVEEVFGA
GSIATLSVLAIVMVAHLVLAEOAPTYLIGRAKTAHYCAPLHWRTSPITIR
AGDIRIAKILISLVGEMTSRTEGAEDPDGSRGDLISGTGPDDEBEVY
GALDIRIPYGDVAIVDEVAAYSTDDLEPTNDRADSPPHRYPLVGESEMDFERGV
YLPEARDDLDLRDGLTALAEVVAHPLVFPADMSDAIDRFQAEQOLAFTVDADTG
GVVGWYITDYLATIGVEEDPTDEAAKSQS"
complement(6599. .6730)
/ gene="VNG0506H"
complement(6599. .6730)
/ gene="VNG0506H"
/ note="hypothetical protein"
/ codon_start=1
/ transl_table=11
/ product="Vng0506h"
/ protein_id="AAG19038.1"
/ db_xref="GI:10580110"
/ translation="MGEPPTPAAGYGVSLMDALSALVFGVYFLTVNAGSAARLD"
complement(6931. .7803)
/ gene="VNG0507C"
complement(6931. .7803)
/ gene="VNG0507C"
/ note="conserved hypothetical protein"
/ codon_start=1
/ transl_table=11
/ product="Vng0507c"
/ protein_id="AAG19039.1"
/ db_xref="GI:10580111"
/ translation="MPMRRLDCGVLPPTIDOLHIPTGTTVEKHDVADSVLIGAS
QURLGVGHNVVAGSERVSDGIEAAGDQRLDMWCVTGTVLVEGSAVYGERHTIER
IYVGDLDIDVDYDIEGFEASGWIYTRDMPRLTFYGNITLHLILGAEADQELVE
ELAADDRCGRPLTIPQASVSDDAWRVSTPATIGDCCRLGNIRATITDVGRCGNVVS
LRAQNDVTVGAGTKIHGDVYTRNGSVHVGODAVVLDVSGHDVTIDEADVDSVIRAR
GEMRVGSVADRDE"
7872. .8069
/ gene="VNG0508H"
7872. .8069
/ gene="VNG0508H"
/ note="hypothetical protein"
/ codon_start=1
/ transl_table=11
/ product="Vng0508h"
/ protein_id="AAG19040.1"
/ db_xref="GI:10580112"
/ translation="MTSLSVADGDVVYEGTEFLERLEKPLIEDATQSDHYVTDHDL
KLVEPNPTLSGGERRIGITLD"
8119. .8256
/ gene="VNG0509H"
8119. .8256
/ gene="VNG0509H"
/ note="hypothetical protein"
/ codon_start=1

```

gene
 5336. .6477
 /gene="cysg"
 /note="represents a fusion of genes with similarity to
 cbix of *Bacillus megaterium* and the 5' region of cysg of
Salmonella typhimurium"
 5336. .6477
 /gene="cysg"
 /note="cobaltochelatase; N-terminal similar to cbix of
Bacillus megaterium; C-terminal similar to
 cobaltochelatase portion of Cysg of *Salmonella*
 typhimurium"
 /codon_start=1
 /transl_table=11
 /product="Cysg"
 /protein_id="AAK67502.1"
 /db_xref="GI:15418796"
 /translation="MTDLVPLVIAHCTDAGLACTRAFADEVRALPGVHELECF
 ELAEDPVAGAVHNAHLHPDAVPSDEPLVLPMLNTGHNVDIPETIEGRDGR
 VSYGGPLDPKRVKLVKEDRINALAPADGPMRADDTSLVLRGALTTNAAEHYR
 LRVYGEVEFGAEPSEIQVVRPSVEALTVADAGATQLVGPFLFRGRLEWYS
 EOVDAWLETHPGIEVRIIDVIGSPPLIAEVFADRYREOYGEVNGDGAAYISGVLA
 GRRVLVAGSHVAERIRPLIRIAGARVHVADNAGTRVRIAGOGVMDQCGFTESD
 VDEWPFVLAASNDPEVNARYSAEAEKQVFCVRSKSDGTATTPTEAGGITVAAY
 GHRNPRSVRYDELTKALQV"
 complement(6509. .7801)
 /gene="cbid"
 complement(6509. .7801)
 /gene="cbid"
 /note="function unknown"
 /codon_start=1
 /transl_table=11
 /product="CbID"
 /protein_id="AAK67503.1"
 /db_xref="GI:15418797"
 /translation="MSAGOPDATGSDAGOPHTHGDSGASAPAGOPPHRYSSDPA
 GQNDRLAREGRROTSGREKOLTSIGLPQGTGACAAARAWSLHTGNFDP
 VEVEPAGRPAPALTEFQIGDGTAAATKAGDDPDVDAVIRATVERPGSGI
 TFRAGSGVGTVPRLPLSGVGEPIAPVPRITYAEMLAADRSVNGSPDYVLTSI
 DGEIOIAQRTAKPKIKIGLISVIGTGVVVPVSCSAMASTHEGIDVARADGASHAA
 ACTGSTSORIAKEEYPDVELLDGDFAGVAVKYLHSHPLRLTICGFFAKMSLNGY
 LDLHSHRVEDQELARLARAREGGDELYAAGAARTASQAYOLSDAGIELGLVAR
 AAKQAATVDAIDIEVICTPRAGITIGRSPTAAH"
 complement(7798. .8451)
 /gene="cbic"
 /note="cobH"
 complement(7798. .8451)
 /gene="cbic"
 /function="catalyzes migration of methyl group from C-11
 to C-12"
 /note="cobyrinic acid synthase"
 /codon_start=1
 /transl_table=11
 /product="Cbic"
 /protein_id="AAK67504.1"
 /db_xref="GI:15418798"
 /translation="MSDIYLRSGSDIYRESFRIIRDEADLSRPDDLEPYAVNTHAA
 DPAIADIATFTGGEAARTALRACAPILCDSSMTATGIIIRSLPNDNVTIKDPR
 LAGIAEEKGVTSAVDLMHEGRLEGSIVAIIGNAPTLFRLIELVHTEGEPAAVI
 GTPGVGAAESKQALVDDACIDYITLIGRGGGSAIAVAINALASSLELTNEHE"
 complement(8548. .9138)
 /gene="cb1t"
 /note="cobL"
 complement(8548. .9138)
 /gene="cb1t"
 /function="decarboxylation of ring C acetate"
 /note="cobalt precorrin 8 decarboxylase"
 /codon_start=1
 /transl_table=11
 /product="Cb1t"
 /protein_id="AAK67505.1"
 /db_xref="GI:15418799"
 /translation="MTTDPDSILGRTPLPESHVPHDGLITKPIRAVALAALRPLP

gene
 9211. .9978
 /gene="cb1j"
 complement(9211. .9978)
 /function="reduction of C-18 to C-19 double bond"
 /note="cobalt dihydroprecorrin 6 synthase"
 /codon_start=1
 /transl_table=11
 /product="Cb1j"
 /protein_id="AAK67506.1"
 /db_xref="GI:15418800"
 /translation="MSTGSIDPEVLVGGTRLARQLARTLTDAGISALTSLAGRRAP
 ROLPGTRHGGEGEGVGLARMLHDNRPRVAVATHAFAATISAAHARACRADPLAR
 LVPSTMAQPDATFTWIVADNASAVRVALPDPVLLTVGRQATARYLAIGDDITHR
 VIDAPDEGPPRRRLNARGPSEAEELMDPGRIRITLVTKDSGGQPAKALVYA
 ARGARVAMTARPPVPDYGIVLHDVADTDMVRSALG"

Query Match 76.3%; Score 20.6; DB 1; Length 11639;
 Best Local Similarity 85.2%; Pred. No. 2.5e+03;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 ggggtcgagctcgagctcgagggggg 27
 ||||| ||||| ||| |||||
 Db 4596 ggggtcgagctcgagctcgagggggg 4622

RESULT 11
 AE005003/c
 LOCUS
 DEFINITION
 AE005003 AE004437
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Halobacterium sp. NRC-1.
 Halobacterium sp. NRC-1.
 Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 Halobacteriaceae; Halobacterium.

REFERENCE
 AUTHORS
 1 (bases 1 to 14379)
 Ng,W.V., Kennedy,S.P., Mahairas,G.G., Bergquist,B., Pan,M.,
 Shukla,H.D., Lasky,S.R., Balliga,N., Thorsson,V., Shproga,J.,
 Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Weir,R., Geo,Y.A.,
 Leitauer,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W.,
 Maddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H.,
 Isenbarger,T.A., Peck,R.F., Pohnschrod,M., Spudich,J.L.,
 Jung,K.H., Alam,M., Freitas,T., Hou,S., Daniels,C.J., Dennis,P.P.,
 Omer,A.D., Ehardt,H., Lowe,T.M., Liang,P., Riley,M., Hood,L. and
 Dassarma,S.
 From the cover: genome sequence of halobacterium species NRC-1
 Proc. Natl. Acad. Sci. USA 97 (22), 12176-12181 (2000)

TITLE
 JOURNAL
 PUBLISHED
 REFERENCE
 AUTHORS

2 (bases 1 to 14379)
 Ng,W.V., Kennedy,S.P., Mahairas,G.G., Bergquist,B., Pan,M.,
 Shukla,H.D., Lasky,S.R., Balliga,N., Thorsson,V., Shproga,J.,
 Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Weir,R., Geo,Y.A.,
 Leitauer,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W.,
 Maddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H.,
 Isenbarger,T.A., Peck,R.F., Pohnschrod,M., Spudich,J.L.,
 Jung,K.H., Alam,M., Freitas,T., Hou,S., Daniels,C.J., Dennis,P.P.,
 Omer,A.D., Ehardt,H., Lowe,T.M., Liang,P., Riley,M., Hood,L. and
 Dassarma,S.
 Direct Submission
 Submitted (14-JUL-2000) Institute for Systems Biology, 4225
 Roosevelt Way NE, Seattle, WA 98105, USA

FEATURES
 source
 1.14379
 /organism="Halobacterium sp. NRC-1"
 /strain="NRC-1"
 /db_xref="taxon:64091"
 79. .825
 /gene="pbd"
 /note="VNC0500C"

[illegible]

FEATURES
source Location/Qualifiers
1. .111188
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub-species="japonica"
/db_xref="taxon:4530"
/chromosome="3"
/clone="OJ1212_C05"

BASE COUNT 32284 a 24371 c 24119 g 30054 t 360 others
ORIGIN

Query Match 77.0%; Score 20.8; DB 2; Length 111188;
Best Local Similarity 91.7%; Pred. No. 1.5e+03;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ggtcgaactcgaactcgaagggg 26
||||| ||||||| |||||||

Db 91111 GGTCCGCGTCGACGACGAGGGG 91088

RESULT 7
AR102311/c 1438 bp DNA linear PAT 14-FEB-2001

LOCUS AR102311 Sequence 26 from patent US 6083902.
DEFINITION AR102311
ACCESSION AR102311
VERSION AR102311.1 GI:12813109
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1438)
AUTHORS Cedarholm-Williams,S.Anthony.
TITLE Recombinant fibrin chains, fibrin and fibrin-homologs
JOURNAL Patent: US 6083902-A 26 04-JUL-2000;
FEATURES
source Location/Qualifiers
1. .1438
/organism="unknown"
BASE COUNT 454 a 293 c 322 g 369 t
ORIGIN

Query Match 76.3%; Score 20.6; DB 6; Length 1438;
Best Local Similarity 85.2%; Pred. No. 3.6e+03;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcgaactcgaactcgaagggg 27
||||| ||||||| |||||||

Db 36 GGGGCCGCGGTGACCTCGAGGGGG 10

RESULT 8
AF340167/c 6219 bp DNA linear BCT 29-MAR-2001

LOCUS AF340167 Streptomyces verticillius polyketide synthase gene, partial cds.
DEFINITION AF340167
ACCESSION AF340167
VERSION AF340167.1 GI:13487280
KEYWORDS
SOURCE Streptomyces verticillius.
ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE 1 (bases 1 to 6219)
AUTHORS Du,L., Sanchez,C., Chen,M., Edwards,D.J. and Shen,B.
TITLE A locus encoding nonribosomal peptide synthetase and polyketide
synthase functions in the bleomycin producer Streptomyces
verticillius ATCC15003
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 6219)
AUTHORS Sanchez,C., Du,L., Chen,M., Edwards,D.J. and Shen,B.
TITLE Direct Submission
JOURNAL Submitted (24-JAN-2001) Chemistry, University of California, One
Shields Avenue, Davis, CA 95616, USA

FEATURES
source Location/Qualifiers
1. .6219
/organism="Streptomyces verticillius"
/strain="ATCC15003"
/db_xref="ATCC:15003"
/db_xref="taxon:29309"
<1. .4085
/note="SVKS"
/codon_start=3
/transl_table=11
/product="polyketide synthase"
/protein_id="AAK27473.1"
/db_xref="GI:13487281"
/translation="SARGPESGAVTGRPVSGEPSTLTGSGVPEPSTATDPVGE
PSVATGATLASEGPEPMATDIAASAASOVAALRADRGHDYPPFVAALKMAASHCAA
ATAFLRPDPAPPTADLLRGVGPVPHATFLSVLPDLRHRRLVGGDSGGATGILPDR
FDYTKRLDILIREPPKAPBGRALHGLRVLRAGGHRPEHLLSGGLADLPOLYDT
GVIGRHNRLAQLNLNVVERPDDRPRLRVLEAGAGGTTALPLPLPERIRYFT
DVSPAFLTRAEHREAAVDELTFETRLDLDVPAAGLPBEGGFVVANALHTARVEA
AVRNVAALAAPNGILLAVESHDEYVLAFLGALDFTMDRDHHERHSPLLTADRNA
LFTRCGETDVVFTGDDTPACRDESVILATAPRRPVPHPYPYPPSPGGRHMTVA
ETRGELPIAREVVAARLPGPDGPAHSVEAVDDVPAHRPSTPSRGATGICILGALPL
EPAAVTITTRRAALRLRITAACPRRDCARFALVYVTPRSSGALPPPERPLADPAV
WGVRKLANEPELDVRTISLDTRGTETDARRLANELTATGETEVALTRSGRRAP
RELPLPAGHTTTTAETHATAVALEIRTPGSLYRPAMWCAEPVGPSEVALDVSAALN
YRDIMQAVGILPEADEGTEETEAEPGIECGVVTAVDGGVTVRPDDRFFALPASIA
SRTYVQAVGRMRPHTMSFEAATLPVVYATVHYSLHDLARLPGETVLVHGAGGVG
LATLRVARGARVIAVAGTPAKRELRLHGLADHVDSRGLDPAHAHVDDTGGGVV
VYNSLGEATTRLELRPGRFEVEGKRDIYENQPLLRPFPHNLAKFQVDTLSAF
DPDQARLFAEVRTRVHDSYRPLPHSAVPAARVAEAFRLQSRHVAKYVTFDEL
EPAVVERPVYAPRRFSQGYLVLTGTSFGAAGARWLADGVRRLLVSRGHPAEA
AGLADLAAAGVATATYAADADPDGMRVIRLIDATGHWVRLVCAHMLDADPLAD
LTPEREAAVLTPEKTGAVALERLLADPLETFLLYSDDTFGLNROAAVYANLHE
ALARORRGELPQOTLAWGVIDETGVVARNDLARTLVAGIEPLSREALTTGTLIS
OGATVAGVGRYMAARLRYVLPASPRCSILPEDELOADTHEELRTIATMPVEA
ARAVAAATLRADLVHLDERNTDHRRLDDEYGLDLSMLAEILVSKRRDIDIPLE
LRSRGTTRDISHLVLRIGITVADDPPEPARDEQSVGER"

BASE COUNT 774 a 2496 c 2173 g 776 t
ORIGIN

Query Match 76.3%; Score 20.6; DB 1; Length 6219;
Best Local Similarity 85.2%; Pred. No. 2.8e+03;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcgaactcgaactcgaagggg 27
||||| ||||||| |||||||

Db 923 GGGTCCGCGGTGACGAGGCG 897

RESULT 9
XU043958 7290 bp DNA linear SYN 28-JAN-1999

LOCUS XU043958 Cloning vector pRCMV-luc luciferase gene, complete cds.
DEFINITION XU043958
ACCESSION U43958
VERSION U43958.1 GI:4097011
KEYWORDS
SOURCE Cloning vector pRCMV-luc.
ORGANISM Cloning vector pRCMV-luc.
REFERENCE 1 (bases 1 to 7290)
AUTHORS Sakamoto,N., Ito,Y., Wu,G.Y. and Wu,C.H.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1995) N. Sakamoto, Division of Gastroenterology,
University of Connecticut Health Center, Farmington, CT 06030, USA

```

gene
  /translation="MDALLPVIYLAGLAAVIGSF7LWASRVRRRIAGACGACALAS
  YEAFRTYAVSHVEIRAQAEERSPILSPDWMWAPADGGAEGAFRTASGSSNSR
  RGRARWRRRRRIATGSHGCHGAGGPGDS"
  complement(736..1902)
  /gene="SCF34.03c"
  /note="SCF34.03c"
  /codon_start=1
  /label="SCF34.03c"
  /product="putative monooxygenase, len: 388 aa;
  similar to SW:YRXX_BACFR (EMBL:M37699), tetr, Bacteroides
  fragilis tetracycline resistance protein from transposon
  Tn4351/Tn4400 (388 aa), fasta scores: opt: 582 z-score:
  637.6 E(): 3.7e-28, 33.9% identity in 351 aa overlap. Also
  similar to monooxygenases e.g. TR:P95555 (EMBL:AB000564)
  Spingomonas sp. salicylate hydroxylase (395 aa) (27.2%
  identity in 360 aa overlap). Similar to others from
  S.coelicolor e.g. TR:Q92416 (EMBL:AF033707) S.coelicolor
  possible salicylate hydroxylase (420 aa) (32.9% identity
  in 340 aa overlap). Contains Pfam match to entry PF01360
  Monooxygenase, Monooxygenase"
  /codon_start=1
  /transl_table=11
  /label="SCF34.03c"
  /product="putative monooxygenase"
  /protein_id="CAB5314.1"
  /db_xref="GI:5763935"
  /translation="MNSPTPAARISIIIGAGPGGLTCARVLRHGIATVYRDDPA
  SRDGGSLDHEEDGQALREAGLLEEFALRCSEQERRRDTGTGLRGLRDPGE
  TAPRERDQGLGLLESIDACTVWRGHLESVSGABEPRLITFDSTVETDLYIG
  ADGARSRRAVSDAVPRITGVGFLEAWDDGSAESELVSGRSAAVACQKELF
  AQRNSGHRVAVMRVALDMWTAGLRPDDVDGIAIRLLAYAGSPRLIMTEND
  GPYVDRPLFALPHTWRTPGVTLLGDAAHMLPPGIVGNLAMDGLALALASA
  TVDDAVRTYEKMLPRSAIACMLEGAGFLLEPDADLARLAGADGGE"
  complement(826..1443)
  /gene="SCF34.03c"
  /note="Pfam match to entry PF01360 Monooxygenase,
  Monooxygenase, score 59.40, E-value 7.7e-14"
  2051..2662
  /gene="SCF34.04"
  /note="SCF34.04"
  /product="putative tetr-family transcriptional
  regulator, len: 203 aa; shows weak similarity to
  SW:TCMR_STRCA (EMBL:M80674), tcmr, Streptomyces.
  glaucoscens tetracycline ctmr transcriptional repressor
  (226 aa), fasta scores: opt: 201 z-score: 254.2 E():
  8.4e-07, 32.2% identity in 152 aa overlap and to many
  putative transcriptional regulators. Contains Pfam match
  to entry PF00440 tetr. Bacterial regulatory proteins, tetr
  family. Contains probable helix-turn-helix motif at aa
  38-59 (Score 1212, +3.31 SD)"
  /codon_start=1
  /transl_table=11
  /label="SCF34.04"
  /product="putative tetr-family transcriptional regulator"
  /protein_id="CAB5315.1"
  /db_xref="GI:5763936"
  /translation="MPDSPTGRRRRKKAATQATADAALQFLHEGDFRVSVDVAV
  RADVSTTLFAHPSKESLYPDQGVDAALAAVREPDGGVGAALRAHLESVPV
  IYDPRRARFTALVSDTPALREYSRRMRHADTISTYIAEELDEPDSDSACALARE
  VLEIPALVTGRDPRATVEYFDLIDHGRSQAGPPPOADR"
  2111..2242
  /gene="SCF34.04"
  /note="Pfam match to entry PF00440 tetr, Bacterial
  regulatory proteins, tetr family, score 45.10, E-value
  8.7e-11"
  87e-11"
  complement(2644..3867)
  /gene="SCF34.05c"
  /complement(2644..3867)
  /gene="SCF34.05c"
  /note="SCF34.05c"
  /codon_start=1
  /transl_table=11
  /label="SCF34.05c"
  /product="putative monooxygenase, len: 407 aa;
  similar to TR:Q92164 (EMBL:AF033534), namh, pseudomonas
  stutzeri salicylate hydroxylase (389 aa), fasta scores:
  opt: 328 z-score: 383.8 E(): 5.1e-14, 27.0% identity in

```

```

RESULT 6
AC092262/c
LOCUS
DEFINITION
AC092262.1 GI:14578165
VERSION
AC092262.1
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzae; Oryza.
1 (bases 1 to 111188)
Buell,R., Hsiao,J., Zismann,V., Moffat,K.M., Hill,J.,
Gansberger,K., Burgess,S., Jarrahl,B., Shvartsbeyn,M., Brenner,M.,
Ciecko,A., Pai,G., Vanaken,S., Hansen,C., Uterbach,T.,
Feldlyum,T., Khalak,H.G., Yuan,Q., Quackenbush,J., White,O.,
Salberg,S., and Fraser,C.
Oryza sativa ssp. japonica cv. Nipponbare O01212_C05 BAC genomic
sequence
Unpublished
2 (bases 1 to 111188)
Buell,R.
Direct Submission
Submitted (30-JUN-2001) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 7495 7538: contig of 7494 bp in length
* 7539 11151: contig of 3613 bp in length
* 11152 11195: gap of unknown length
* 11196 21418: contig of 10223 bp in length
* 21419 21462: gap of unknown length
* 21463 52544: contig of 31082 bp in length
* 52545 52588: gap of unknown length
* 52589 60154: contig of 7566 bp in length
* 60155 60198: gap of unknown length
* 60199 69854: contig of 9656 bp in length
* 69855 69898: gap of unknown length
* 69899 79565: contig of 9667 bp in length
* 79566 79608: gap of unknown length
* 79609 90877: contig of 11269 bp in length
* 90878 90920: gap of unknown length
* 90921 111188: contig of 20268 bp in length.

```

403 aa overlap and to the N-terminal half of some epoxidases e.g. SW:ABA2_LYCES (EMBL:Z83835) Lycopersicon esculentum zeaxanthin epoxidase precursor (663 aa) (26.8% identity in 336 aa overlap). Also similar to many putative oxidoreductases e.g. TR:O54177 (EMBL:AF021411) Streptomyces coelicolor possible oxidoreductase (397 aa) (56.1% identity in 403 aa overlap). Contains Pfam match to entry PF01360 Monooxygenase, Monooxygenase"

Query Match 77.0%; Score 20.8; DB 1; Length 38995;
Best local Similarity 91.7%; Pred. No. 1.8e+03;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ggtcgcagtcgacgtcga999999 26
Db 35414 GGTGCGAGTCGCGTCGAGGTCG 35391

```

/codon_start=1
/transl_table=11
/product="delta 4, 5-alpha steroid dehydrogenase"
/protein_id="AA080517.1"
/db_xref="GI:404686"
/translation="MSNVKHYSTIMPGEVLVDGSDAEVONSADSDVYVVGNGCAGA
SAIAIEREGGAELVIERFSGGASVLSGGVYAGVAPTRRRPARFEAMATYAKH
EVNGVSEDETLARFSRDSVTNLNMLEKOGATEPMGYKTSPADGMVLYSGNEV
PAYNGPOLTKKPPRGHRYVAKGSGAMFEALOKSTLAHGAATTLQARVORLVREK
SGRVLGEVYVLPEDGPRTERHKIDELVAKSACIRRVPRVAVVRRSRRARSASA
TSVPKWCPCPLAITSINCSMRRYKRGMLTGAGCGSGRLGOSVIGIADLNN
ISAMRITPPSPKGLVYNICEFCNEQVGAOKLGTETMEKOGQAWLIDISNVR
QAMWCFEGGLAFQSPALALMYKVAIKGSVDLAKLRMDAAVLAQLQPRANPA
RGEIEDPLKSDMDMREHFRGSLFALDISISQKMPFLAVLSIGELKVEDNGAVIDGA
GYDIPGLVAGVPLVWLPVRYT"

BASE COUNT      417 a      602 c      751 g      425 t
ORIGIN
Query Match      77.0%; Score 20.8; DB 1; Length 2195;
Best Local Similarity 91.7%; Pred. No. 2.9e+03;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 gtcgaagtcgagtcgagggggg 27
      ||||| ||||| ||||| |||||
Db 24 gtcgacctcgacctcgaggggg 1

RESULT 5
LOCUS      SCF34/c      38995 bp      DNA      linear      BCF 24-SEP-1999
DEFINITION Streptomyces coelicolor cosmid F34.
ACCESSION  AL109974
VERSION     AL109974.1 GI:5763932
KEYWORDS   ABC-transporter; ATP-binding; chig; chitinase; DEAD-box; esterase;
            formyltransferase; iron; lipoprotein; Murr-family; monooxygenase;
            peptide synthetase; permease; RNA helicase; siderophore;
            tetr-family; transcriptional regulator.
SOURCE     Streptomyces coelicolor A3(2)
ORGANISM   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
REFERENCE  1 (bases 1 to 38995)
            Redenbach,M., Kleser,H.M., Denapalte,D., Eicher,A., Cullum,J.,
            Kinashi,H. and Hopwood,D.A.
            A set of ordered cosmids and a detailed genetic and physical map
            for the 8 Mb Streptomyces coelicolor A3(2) chromosome
            Mol. Microbiol. 21 (1), 77-96 (1996)
            97000351
JOURNAL    MEDLINE
REFERENCE  2 (bases 1 to 38995)
            Saunders,D.C. and Harris,D.
            unpublished
JOURNAL    REFERENCE
            3 (bases 1 to 38995)
            James,K.D., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
            Direct Submission
            Submitted (24-SEP-1999) Streptomyces coelicolor sequencing project,
            Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
            CB10 1SA E-mail: barrellesanger.ac.uk Cosmids supplied by Prof.
            David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
            Colney, Norwich, Norfolk NR4 7UH, UK
            Notes:
            Streptomyces coelicolor sequencing at The Sanger Centre is funded
            by the BBSRC and Beowulf Genomics
            Details of S. coelicolor sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL: http://www.sanger.ac.uk/Projects/S.coelicolor/)
            CDS are numbered using the following system eg SC7B7.01c. SC (S.
            coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
            strand).
            The more significant matches with motifs in the PROSITE database
            are also included but some of these may be fortuitous.
            The length in codons is given for each CDS.
            Usually the highest scoring match found by fasta -o is given for

```

```

CDS which show significant similarity to other CDS in the database.
The position of possible ribosome binding site sequences are given
where these have been used to deduce the initiation codon.
Gene prediction is based on positional base preference in codons
using a specially developed Hidden Markov Model (Krogh et al.,
Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot
program of Bibb et al., Gene 30:157-66(1984) as implemented at
http://www.nih.gov.jp/
jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the
correct initiation codon. Where possible we choose an initiation
codon (atg, gty, ttg or (atc)) which is preceded by an upstream
ribosome binding site sequence (optimally 5-13bp before the
initiation codon). If this cannot be identified we choose the most
upstream initiation codon.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions.
Cosmid F34 Overlaps with cosmid F80 on the AseI-F genomic
restriction fragment.
FEATURES
    source
        1..38995
            /organism="Streptomyces coelicolor A3(2)"
            /strain="A3(2)"
            /db_xref="taxon:100226"
            /clone="cosmid F34"
            1..527
            /note="sequence corresponding to Streptomyces coelicolor
            gene for Chig, partial cds (EMBL:AB017013) from 1449 to
            1975"
            1..101
            /gene="SCF34.01"
            /note="overlap with Streptomyces coelicolor cosmid F80
            from 10824 to 10924"
            <1..273
            /gene="SCF34.01"
            /note="SCF34.01, chig, chitinase, partial CDS, len: >90
            aa; identical to the C-terminus of TR:0929W4
            (EMBL:AB017013), chig, Streptomyces coelicolor chitinase
            (fragment) (244 aa), fasta scores: opt: 638 z-score: 812.8
            E(): 0, 100.0% identity in 90 aa overlap. Highly similar
            to SC5H1.29c (EMBL:AL049863), chIF, Streptomyces
            coelicolor chitinase (296 aa) (81.1% identity in 90 aa
            overlap). Contains Pfam match to entry PF00182
            Glyco_hydro.19, Chitinases class I"
            /codon_start=1
            /transl_table=11
            /label="chig"
            /product="chitinase"
            /protein_id="CAB53312.1"
            /db_xref="GI:5763933"
            /translation="DLNNPDLYQNSAVAKTGLMYNTQRCPTMTPHDAMVNGAG
            FGFTIRISNGLECDGNGPQVOSRIDNTERFOLLGVEPGNLSC"
            1..273
            /gene="SCF34.01"
            /note="chig"
            1..270
            /gene="SCF34.01"
            /note="Pfam match to entry PF00182 Glyco_hydro.19,
            chitinases class I, score 87.60, E-value 7.6e-24"
            complement(270..671)
            /gene="SCF34.02c"
            complement(270..671)
            /gene="SCF34.02c"
            /note="SCF34.02c, hypothetical protein, len: 133 aa;
            unknown function, possible CDS suggested by GC frameplot.
            contains an Arg-rich region"
            /codon_start=1
            /transl_table=11
            /label="SCF34.02c"
            /product="hypothetical protein"
            /protein_id="CAB53313.1"
            /db_xref="GI:5763934"

```



```

misc_feature
/db_xref="GI:10303264"
/translation="MNDRMVWIDCEWNTGLSDDLIEVAALVTSENLILGCVDI
IRPERALTEMEVVEVREHMTASGLAEDGTTLADAEAOVLAYREHKEGKAPIC
GNSVGTDRGFTLRDMATLEGGYLAHYRVDSSTIKELARWRPRAYFNSPKNGHRLA
DIRESIAELRYREAVFVFPQPPDSDTAAIAAKHVSAG"
123. .641
/gene="orfa"
/notes="Pfam match to entry PF00929 Exonuclease,
Exonuclease, score 76.80, E-value 4.6e-19"
835. .907
/notes="tRNA His anticodon GTG, Cove score 71.04"
/product="tRNA-His"
914. .951
/notes="possible stem loop. Score 57: 18/18 (100%) matches,
0 gaps"
complement(1275. .2330)
/gene="2SCC13.02c"
complement(1275. .2330)
/gene="2SCC13.02c"
/notes="2SCC13.02c, lacI-family transcriptional regulatory
protein, len: 351 aa: highly similar to TR:Q9X9R3
(EMBL:AJ009798) Streptomyces reticuli CebR protein, 350
aa; fasta scores: opt: 2011 z-score: 2174.4 E(): 0; 89.4%
identity in 350 aa overlap and C-terminal region identical
to TR:BA03462 (EMBL:AB036424) Streptomyces coelicolor
transcriptional regulator (fragment) ORF5-H, 193 aa.
Contains Pfam matches to entries PF00356 lacI, Bacterial
regulatory proteins, lacI family and PF00532
Peripla_BP_1like, Periplasmic binding proteins and sugar
binding domain of the lacI family and match to Prosite
entry PS00356 Bacterial regulatory proteins, lacI family
signature. Also contains a possible helix-turn-helix motif
at residues 14. .35 (+6.05 SD)"
/codon_start=1
/transl_table=11
/product="lacI-family transcriptional regulatory protein"
/protein_id="CAC10103.1"
/db_xref="GI:10303265"
/translation="MASHGVRGSGRPILIEVAAAGVGTGVSRTVINGSPRVADT
RAVVAALVETGVPNTAARALAAANTDAIALVPEPEPRFAEYFSDMLKGVSEL
SETEMQLLITFGSPDRERRLAQYTLAAHRYDVGLVSHADPLDGLDITLPIVVIS
GPRSAEPLIASVSDNYGARSVAEHLISRGKRVAAHTTGAAVGAQRVRYGDAAL
REAGHEVEGLIEPDEFTTEGGRAMAEILRRHPVDVAFADVTAGAGCVLRDAG
RRIPDVALVGYDDSDAIAIRHMEPPLTVSVOPIEEMGRAMIDLLIEIADRRPAASRL
ERHVVATLELVERRS"
complement(1386. .2123)
/gene="2SCC13.02c"
/notes="Pfam match to entry PF00532 Peripla_BP_1like,
Periplasmic binding proteins and sugar binding domain pf
the lacI family, score 159.90, E-value 4.3e-44"
complement(2214. .2297)
/gene="2SCC13.02c"
/notes="Pfam match to entry PF00356 lacI, Bacterial
regulatory proteins, lacI family, score 35.50, E-value
1.5e-08"
complement(2229. .2285)
/gene="2SCC13.02c"
/notes="PS00356 Bacterial regulatory proteins, lacI family
signature"
complement(2334. .2337)
2685. .2690
2696. .4021
/gene="2SCC13.03"
2696. .4021
/gene="2SCC13.03"
/notes="2SCC13.03, probable sugar binding secreted protein,
len: 441 aa: highly similar to TR:Q9X9R7 (EMBL:AJ009797)
Streptomyces reticuli Cbp protein precursor (fragment)
CebR, 444 aa; fasta scores: opt: 1490 z-score: 1672.5 E():
0; 51.2% identity in 445 aa overlap and similar to
SW:LACE_AGRAD (EMBL:X66596) Agrobacterium radiobacter
lactose-binding protein precursor lacE, 422 aa; fasta
scores: opt: 263 z-score: 298.4 E(): 4.1e-09; 23.1%

```

```

misc_feature
/translation="MKAARRGSARRVYMAIASLGAGILLACDAGDDESSSGDS
SGRTITLITLFTMGFKENGALDYDEKLNPDINIEVTERENYTPALVNLITNSG
LODVOAIEVGNIAEVVAQADKEEDMSKAQVAKMDMKQAQATTKGATIGLGTD
IGPMALVCKDLFEKAGLPTDREEVSKIMAGMKNFTEAGKYGAGKDYFENDSPG
GLINALSDEDEKFTDASGVYIKTNPAVKDFDITFAEAGLVQSQTQPPANDOT
ISNSLFAIVACPPMMLGTTKAKSQPSDSACKWVQAPAKGNGGTFLLGPKSGKHAKE
AOKLVWTLAPEDQAKLPTQKSGSPSAPAAVYLPVOTGKNDMTGDAPIGELFAKAE
QIPTVGPKDDIVQOGLTNDGVILVTOGKSAEDAMDNAVKTIDNNLEK"
2750. .2782
/gene="2SCC13.03"
/notes="2SCC13.03 Prokaryotic membrane lipoprotein lipid
attachment site"
3146. .4006
/gene="2SCC13.03"
/notes="Pfam match to entry PF01547 SBP bacterial_1,
Bacterial extracellular solute-binding protein, score
35.20, E-value 7.1e-09"
4012. .4015
/gene="2SCC13.03"
4027. .5040
/gene="2SCC13.04"
4027. .5040
/gene="2SCC13.04"
/notes="2SCC13.04, probable cellulose transport permease,
len: 337 aa: highly similar to TR:Q9X9R6 (EMBL:AJ009797)

```

Query Match 80.7%; Score 21.8; DB 1; Length 20812;
Best local similarity 92.0%; Pred. No. 9, 8e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Db 8192 GTCGACTTCGACGTCGAGGCGG 8168
QY 3 gtcgcagtcgcagtcgcagggg 27
||||| ||||||| |||
8192 GTCGACTTCGACGTCGAGGCGG 8168

```

RESULT 4
CMAD45AHD/C 2195 bp DNA linear BCT 20-SEP-1996
LOCUS CMAD45AHD
DEFINITION Comamonas testosteroni delta 4, 5-alpha steroid dehydrogenase gene,
complete cds.
ACCESSION L23428
VERSION L23428.1 GI:404685
KEYWORDS delta 4, 5-alpha steroid dehydrogenase.
SOURCE Comamonas testosteroni
ORGANISM Comamonas testosteroni
Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
Comamonas.
REFERENCE 1 (bases 1 to 2195)
AUTHORS Florin,C., Kohler,T., Grandguillot,M. and Plesiat,P.
TITLES Comamonas testosteroni 3-ketosteroid-delta(5alpha)-dehydrogenase:
gene and protein characterization
JOURNAL J. Bacteriol. 178 (11), 3322-3330 (1996)
MEDLINE 96236051
FEATURES
source location/Qualifiers
1. .2195
/organism="Comamonas testosteroni"
/db_xref="taxon:285"
/feature_id="ATCC 17410"
89. .94
102. .1694
/EC_number="1.3.99.4"
/notes="3-oxosteroid delta 4(5 alpha)-dehydrogenase"

Query Match 100.0%; Score 27; DB 6; Length 27;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ggggtcagctgcagctcgaaggagg 27
 |||||
 Db 1 GGGGTGCGAGCTGCGAGGAGGGGG 27

RESULT 2
 AX105138 27 bp DNA linear PAT 30-APR-2001
 LOCUS Sequence 36 from Patent WO0122990.
 DEFINITION AX105138
 ACCESSION AX105138
 VERSION AX105138.1 GI:13921288
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 27)
 AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
 TITLE Methods related to immunostimulatory nucleic acid-induced
 JOURNAL Interferon
 Patent: WO 0122990-A 36 05-APR-2001;
 Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
 FOUNDATION (US)

FEATURES
 source location/Qualifiers
 1..27
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Synthetic Oligonucleotide"
 misc_feature 1..2
 /note="Backbone has phosphorothioate linkages."
 misc_feature 3..21
 /note="Backbone has phosphodiester linkages."
 misc_feature 22..26
 /note="Backbone has phosphorothioate linkages."
 misc_feature 27
 /note="Backbone has phosphodiester linkages."
 BASE COUNT 3 a 5 c 16 g 3 t
 ORIGIN

Query Match 100.0%; Score 27; DB 6; Length 27;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ggggtcagctgcagctcgaaggagg 27
 |||||
 Db 1 GGGGTGCGAGCTGCGAGGAGGGGG 27

RESULT 3
 SC2C13 20812 bp DNA linear BCT 23-SEP-2000
 LOCUS Streptomyces coelicolor cosmid 2C13.
 DEFINITION AL442165
 ACCESSION AL442165
 VERSION AL442165.1 GI:10303263
 KEYWORDS
 cellulose hydrolase; cellobiose transport permease; integral
 membrane protein; lact-family transcriptional regulatory protein;
 mutase; orna; secreted protein; secreted sugar hydrolase;
 transcriptional regulatory protein; two component system histidine
 kinase; two component system response regulator.
 Streptomyces coelicolor A3(2).
 SOURCE Streptomyces coelicolor A3(2).
 ORGANISM Streptomyces coelicolor A3(2).
 Bacteria; Firmicutes; Actinobacteria; Actinobacteriales;
 Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 1 (bases 1 to 20812)
 Redenbach M., Kleser, H.M., Denapate, D., Eichner, A., Cullum, J.,
 Kinashi, H. and Hopwood, D.A.
 A set of ordered cosmids and a detailed genetic and physical map
 for the 8 Mb Streptomyces coelicolor A3(2) chromosome

JOURNAL
 MEDLINE
 REFERENCE 97000351
 AUTHORS Seeger,K.J. and Harris,D.
 JOURNAL unpublished
 3 (bases 1 to 20812)
 REFERENCE Cerdeno,A.M., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
 TITLE Direct Submission
 JOURNAL Submitted (22-SEP-2000) Streptomyces coelicolor sequencing project,
 Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridg
 CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
 David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
 Colney, Norwich, Norfolk NR4 7UH, UK
 Notes:
 Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC and Beowulf Genomics
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: <http://www.sanger.ac.uk/projects/s-coelicolor/>)
 CDS are numbered using the following system eg SC787.01c. SC (S.
 coelicolor), 787 (cosmid name), .01 (first CDS), c (complementary
 strand).
 The more significant matches with motifs in the PROSITE database
 are also included but some of these may be fortuitous.
 The length in codons is given for each CDS.
 Usually the highest scoring match found by fasta -o is given for
 CDS which show significant similarity to other CDS in the database.
 The position of possible ribosome binding site sequences are given
 where these have been used to deduce the initiation codon.
 Gene prediction is based on positional base preference in codons
 using a specially developed Hidden Markov Model (Krogh et al.,
 Nucleic Acids Research, 22(22):4768-4778(1994)) and the Frameplot
 program of Bibb et al., Gene 30:157-66(1984) as implemented at
<http://www.nih.gov/jp/jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the
 correct initiation codon. Where possible we choose an initiation
 codon (atg, gtg, ttg or (att)) which is preceded by an upstream
 ribosome binding site sequence (Optimally 5-13bp before the
 initiation codon). If this cannot be identified we choose the most
 upstream initiation codon.
 IMPORTANT: This sequence MAY NOT be the entire insert of the
 sequenced clone. It may be shorter because we only sequence
 overlapping sections once, or longer, because we arrange for a
 small overlap between neighbouring submissions.
 Cosmid 2C13.

FEATURES
 source location/Qualifiers
 1..20812
 /organism="Streptomyces coelicolor A3(2)"
 /strain="A3(2)"
 /db_xref="taxon:100226"
 /clone="cosmid 2C13"
 1..1858
 /note="Previously sequenced DNA fragment. EMBL:AB036424
 Streptomyces coelicolor A3(2) adpA, orna, orf5-h genes for
 adpA homolog, oligoribonuclease, transcriptional
 regulator, partial and complete cds."
 1..120
 /note="nominal overlap with Streptomyces coelicolor cosmid
 SCC105"
 111..713
 /gene="orna"
 /note="2SCC13.01"
 111..713
 /gene="orna"
 /note="2SCC13.01, orna, oligoribonuclease, len: 200 aa;
 identical to previously sequenced TR:BA003461
 (EMBL:AB036424) Streptomyces coelicolor oligoribonuclease
 orna, 200 aa. Contains Pfam match to entry PF00929
 Exonuclease, Exonuclease"
 /codon_start=1
 /transl_table=11
 /product="Oligoribonuclease"
 /protein_id="CAC10102.1"

gene
 misc_feature
 CDS

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:58:46 ; Search time 2778.35 seconds
(without alignments) : 203.364 Million cell updates/sec

Title: US-09-672-126-36

Perfect score: 27
Sequence: 1 ggggagcagtcgacgcagcagggggg 27

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenBank:1:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	27	100.0	27	6	AX104885	AX104885 Sequence
2	27	100.0	27	6	AX105138	AX105138 Sequence
3	21.8	80.7	20812	1	SC2C13	AL442165 Streptomy
4	20.8	77.0	2195	1	CMAD45AHYD	L23428 Commomonas t
5	20.8	77.0	38995	1	SCF34	AL109974 Streptomy
6	20.8	77.0	111188	2	AC092262	AC092262 Oryza sat
7	20.6	76.3	1438	6	AR102311	AR102311 Sequence
8	20.6	76.3	6219	1	AF340167	AF340167 Streptomy
9	20.6	76.3	7290	12	XX043958	U43958 Cloning vec
10	20.6	76.3	11639	1	AE003236	AY033236 Propionib
11	20.6	76.3	14379	1	AE005003	AE005003 Halobacte
12	20.6	76.3	31624	1	SCD63	AL161755 Streptomy
13	20.6	76.3	41944	3	AC005929	AC005929 Leishman
14	20.6	76.3	115873	2	AC017383	AC017383 Drosophi
15	20.6	76.3	135295	8	AP003282	AP003282 Oryza sat
16	20.6	76.3	142268	8	AP003018	AP003018 Oryza sat
17	20.6	76.3	146638	2	AP003252	AP003252 Oryza sat
18	20.6	76.3	154084	8	AP003734	AP003734 Oryza sat
19	20.6	76.3	154248	8	AP003631	AP003631 Oryza sat
20	20.6	76.3	159749	8	AP003020	AP003020 Oryza sat
21	20.6	76.3	175681	3	AC007417	AC007417 Drosophi
22	20.6	76.3	185932	2	AP003714	AP003714 Oryza sat
23	20.6	76.3	190574	2	AC007352	AC007352 Drosophi
24	20.6	76.3	193119	8	AC025907	AC025907 Oryza sat
25	20.6	76.3	215241	8	AF459639	AF459639 Trilicium
26	20.6	76.3	229896	14	AF232689	AF232689 Rat cytom
27	20.6	76.3	261846	3	AE003830	AE003830 Drosophi
28	20.4	75.6	1183	3	DROADHPSGB	L26040 Drosophi
29	20.4	75.6	3184	1	STU09309	U09309 Salimonia
30	20.4	75.6	123953	2	AP003747	AP003747 Oryza sat
31	20.2	74.8	7317	6	AX277884	AX277884 Sequence
32	20.2	74.8	7317	6	AX323559	AX323559 Sequence
33	20.2	74.8	38404	1	SC2G5	AL035478 Streptomy
34	20.2	74.8	120126	2	AC087096	AC087096 Oryza sat
35	20.2	74.8	142711	8	AC087181	AC087181 Oryza sat
36	20.2	74.8	152763	2	AC093713	AC093713 Oryza sat
37	19.8	73.3	10339	12	U02448	U02448 Cloning vec
38	19.8	73.3	12046	1	AE005017	AE005017 Halobacte
39	19.8	73.3	41622	1	SCD25	AL118514 Streptomy
40	19.8	73.3	56917	1	AME16952	Y16952 Amycolatops
41	19.8	73.3	112721	8	AC016780	AC016780 Genomic S
42	19.8	73.3	142737	8	AC027658	AC027658 Oryza sat
43	19.6	72.6	228	8	AY022072	AY022072 Oryza sat
44	19.6	72.6	13856	1	AE004026	AE004026 Xylella f
45	19.6	72.6	14878	1	AE005849	AE005849 Caulobact

ALIGNMENTS

RESULT 1
AX104885
LOCUS AX104885 27 bp DNA
DEFINITION Sequence 1077 from Patent WO0122972.
ACCESSION AX104885
VERSION AX104885.1 GI:13921082
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 27)
AUTHORS Kriegl,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 1077 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)

FEATURES
source
1..27 Location/Qualifiers
BASE COUNT 3 a 5 c 16 g 3 t

ORGANISM
"organism="synthetic construct"
/db_xref="taxon:32630"

THIS PAGE BLANK (UPTO)

RESULT 15

US-08-342-411A-1

; Sequence 1, Application US/08342411A

; Patent No. 5639616

; GENERAL INFORMATION:

; APPLICANT: LIAO, Shutsung

; APPLICANT: SONG, Ching

; TITLE OF INVENTION: UBIQUITOUS NUCLEAR RECEPTOR:

; TITLE OF INVENTION: COMPOSITIONS AND METHODS

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: TX

; COUNTRY: USA

; ZIP: 77210-4433

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/342,411A

; FILING DATE: 18-NOV-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: KITCHELL, BARBARA S.

; REGISTRATION NUMBER: 33,928

; REFERENCE/DOCKET NUMBER: ARCD154

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000

; TELEFAX: (713) 789-2679

; TELEX: 79-0924

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1898 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 71..1450

US-08-342-411A-1

Query Match 76.0%; Score 15.2; DB 1; Length 1898;

Best Local Similarity 85.0%; Pred. No. 1.3e+02;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggtcgtcgagggggggg 20

DB 1437 GGGACGTGACGAGTGAGGG 1456

Search completed: August 10, 2002, 03:06:33
Job time: 16059 sec

APPLICATION NUMBER: 08/202,044
FILING DATE: 23-Feb-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 96,137-A (11274/02148)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1403 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 101..949
US-08-751-344B-1

Query Match 76.0%; Score 15.2; DB 4; Length 1403;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggtcgtcgcagagggggg 20
||||| ||| ||| ||| |||
DB 330 GGGTCATCGCGAGCGGGG 311

RESULT 13
US-08-583-672-1/C
Sequence 1, Application US/08583672
Patent No. 5741673
GENERAL INFORMATION:
APPLICANT: Montlun, Marc R.
TITLE OF INVENTION: A NOVEL HOMEOBOX FACTOR THAT STIMULATES
TITLE OF INVENTION: INSULIN EXPRESSION IN PANCREATIC ISLET CELLS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/583,672
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/106,936
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9422
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1614 base pairs
TYPE: nucleic acid

STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 331..1182
OTHER INFORMATION: /product= "ITF-1 Homeobox-Lyte
OTHER INFORMATION: transcription factor"
US-08-583-672-1

Query Match 76.0%; Score 15.2; DB 1; Length 1614;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggtcgtcgcagagggggg 20
||||| ||| ||| ||| |||
DB 560 GGGTCATCGCGAGCGGGG 541

RESULT 14
US-08-442-884-2
Sequence 2, Application US/08442884
Patent No. 5637490

GENERAL INFORMATION:
APPLICANT: Mutsaers SANO et al.
TITLE OF INVENTION: ALPHA-1,3/4-FUCOSIDASE GENE
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,884
FILING DATE: May 17, 1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:

INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1689 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
US-08-442-884-2

Query Match 76.0%; Score 15.2; DB 1; Length 1689;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggtcgtcgcagagggggg 20
||||| ||| ||| ||| |||
DB 1031 GGGTCGACAGAGCGGGG 1050

RESULT 10
US-08-440-846-1
; Sequence 1, Application US/08440846
; Patent No. 5690939
; GENERAL INFORMATION:
; APPLICANT: Morgan, Robin Wilson
; APPLICANT: Claessens, Johannes Antonius Joseph
; APPLICANT: Sondermeijer, Paulus Jacobus Antonius
; TITLE OF INVENTION: Recombinant vaccine against Marek's
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Organon Teknika Corporation
; STREET: 1330-A Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS-DOS 3.3
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,846
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/699,467
; FILING DATE: 14-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bobrowicz, Donna
; REGISTRATION NUMBER: 32196
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 975 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marek's Disease Virus
; STRAIN: Georgia (Ga)
; IMMEDIATE SOURCE:
; CLONE: PMD11
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 46..918
; OTHER INFORMATION: /label= MD06_antigen
US-08-440-846-1

Query Match 76.0%; Score 15.2; DB 1; Length 975;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggctgcgcagcagggggg 20
||||| ||||||| |||
Db 134 GGGTCGCCGACGAGCAGG 153

RESULT 11
US-08-202-044-1/c
; Sequence 1, Application US/08202044
; Patent No. 5858973
; GENERAL INFORMATION:
; APPLICANT: Habener M.D., Joel F.
; APPLICANT: Miller Ph.D., Christopher P.
; TITLE OF INVENTION: NOVEL TRANSCRIPTION FACTOR AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/202,044
FILING DATE: 23-FEB-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Williams Ph.D., Kathleen A.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: MGH-124XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1403 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 101..949
US-08-202-044-1

Query Match 76.0%; Score 15.2; DB 2; Length 1403;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggctgcgcagcagggggg 20
||||| ||||||| |||
Db 330 GGGTCATCGCGCAGCGGGG 311

RESULT 12
US-08-751-344B-1/c
; Sequence 1, Application US/08751344B
; Patent No. 6210960
; GENERAL INFORMATION:
; APPLICANT: Habener M.D., Joel F.
; APPLICANT: Miller Ph.D., Christopher P.
; TITLE OF INVENTION: NOVEL TRANSCRIPTION FACTOR AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Wilcoff, Ltd.
; STREET: One Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,344B
; FILING DATE: 19-NO. 6210960-1996
; PRIOR APPLICATION DATA:

STATE: New Jersey
COUNTRY: United States
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/233,009
FILING DATE: 25-APR-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Benven, Gerard H
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: 19219
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3901
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-233-009-4

Query Match 76.0%; Score 15.2; DB 1; Length 60;
Best Local Similarity 85.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggctcgtcagagggggg 20
||||| ||| ||| ||| |||
DB 24 GGGTCCTGTCGACGGGGG 43

RESULT 8
US-08-560-231-4
Sequence 4, Application US/08560231
Patent No. 5817760
GENERAL INFORMATION:
APPLICANT: Jacobson, Marlene A
APPLICANT: Johnson, Robert G
APPLICANT: Luneau, Christopher J
APPLICANT: Salvatore, Christopher A
TITLE OF INVENTION: Human Adenosine Receptors
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: NJ
COUNTRY: United States
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh IIci
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,231
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Meredith, Roy D.
REGISTRATION NUMBER: 30,777
REFERENCE/DOCKET NUMBER: 186991A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4678

TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-560-231-4

Query Match 76.0%; Score 15.2; DB 1; Length 60;
Best Local Similarity 85.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggctcgtcagagggggg 20
||||| ||| ||| ||| |||
DB 24 GGGTCCTGTCGACGGGGG 43

RESULT 9
US-09-080-704A-4
Sequence 4, Application US/09080704A
Patent No. 6166181
GENERAL INFORMATION:
APPLICANT: Jacobson, Marlene A
APPLICANT: Johnson, Robert G
APPLICANT: Luneau, Christopher J
APPLICANT: Salvatore, Christopher A
TITLE OF INVENTION: Human Adenosine Receptors
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: NJ
COUNTRY: United States
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/080,704A
FILING DATE: 18 May 1998
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Pat, Richard S.
REGISTRATION NUMBER: 32,586
REFERENCE/DOCKET NUMBER: 18699DB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (732) 594-4958
TELEFAX: (732) 594-4720
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-080-704A-4

Query Match 76.0%; Score 15.2; DB 4; Length 60;
Best Local Similarity 85.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggctcgtcagagggggg 20
||||| ||| ||| ||| |||
DB 24 GGGTCCTGTCGACGGGGG 43

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 5496
TYPE: DNA
ORGANISM: Fungus
US-09-462-284-1

Query Match 79.0%; Score 15.8; DB 4; Length 5496;
Best Local Similarity 89.5%; Pred. No. 69;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggcctgcgcagcagggggg 19
|||||
Db 3225 GTCCTGCGACGACGGGG 3207

RESULT 5
US-09-194-905-7/C
Sequence 7, Application US/09194905
Patent No. 6306627
GENERAL INFORMATION:
APPLICANT: DECKER, Heinrich
TITLE OF INVENTION: ISOLATION OF THE BIOSYNTHESIS GENES FOR
PSEUDO-OLIGOSACCHARIDES FROM STREPTOMYCES GLAUCESCENS
TITLE OF INVENTION: GLA.O AND THEIR USE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/194,905
FILING DATE: 29-JUL-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP97/02826
FILING DATE: 30-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 19622783.6
FILING DATE: 07-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Granados, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 026083/0193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 6854 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-194-905-7

Query Match 79.0%; Score 15.8; DB 4; Length 6854;
Best Local Similarity 89.5%; Pred. No. 68;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 ggcctgcgcagcagggggg 20
|||||
Db 43 GTCCTGCGACGACGGGGG 25

RESULT 6
US-08-349-696-4
Sequence 4, Application US/08349696
Patent No. 5599671

GENERAL INFORMATION:
APPLICANT: Jacobson, Marlene A
APPLICANT: Johnson, Robert G
APPLICANT: Luneau, Christopher J
APPLICANT: Salvatore, Christopher A
TITLE OF INVENTION: Human Adenosine Receptors
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: NJ
COUNTRY: United States
ZIP: 07065

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh IIfx
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/349,696
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: us/08/005945

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Meredith, Roy D.
REGISTRATION NUMBER: 30,777
REFERENCE/DOCKET NUMBER: 186991A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)594-4678
TELEFAX: (908)594-4720
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-349-696-4

Query Match 76.0%; Score 15.2; DB 1; Length 60;
Best Local Similarity 85.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggcctgcgcagcagggggg 20
|||||
Db 24 GGCCTGCTGTCGACGGGGG 43

RESULT 7
US-08-233-009-4
Sequence 4, Application US/08233009
Patent No. 5646156
GENERAL INFORMATION:
APPLICANT: Jacobson, Marlene A
APPLICANT: Johnson, Robert G
APPLICANT: Salvatore, Christopher A
APPLICANT: Salvatore, Christopher A
TITLE OF INVENTION: INHIBITION OF EOSINOPHIL
ACTIVATION THROUGH A3 ADENOSINE RECEPTOR ANTAGONISM
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 99gtcgtcagcagggggg 20
|||||
Db 249 GGGTCGTGCGAGGGGGG 230

RESULT 2
US-08-881-450A-23/c
Sequence 23, Application US/08881450A
Patent No. 6274310
GENERAL INFORMATION:
APPLICANT: Habener, J.F. and Stoffers, D.A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: PANCREATIC DISEASE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESS: Banner & Witcoff, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/881,450A
FILING DATE: June 24, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 11275/7823
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 5658 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
FEATURE:
NAME/KEY: IP11 gene; contig 2.
FEATURE:
NAME/KEY: transcriptional start
LOCATION: nucleotide 2002
FEATURE:
NAME/KEY: translational start codon
LOCATION: nucleotides 2106 through 2108
FEATURE:
NAME/KEY: first coding region
LOCATION: nucleotides 2106 through 2511
FEATURE:
NAME/KEY: intron 1
LOCATION: nucleotides 2512 through 5858
US-08-881-450A-23

Query Match 92.0%; Score 18.4; DB 4; Length 5658;
Best Local Similarity 95.0%; Pred. No. 5.5;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 99gtcgtcagcagggggg 20
|||||
Db 2135 GGGTCGTGCGAGGGGGG 2116

RESULT 3
US-08-762-308-10/c
Sequence 10, Application US/08762308
Patent No. 5925548
GENERAL INFORMATION:
APPLICANT: Beutler, Bruce A.
TITLE OF INVENTION: MODIFIED RECEPTORS THAT CONTINUOUSLY
TITLE OF INVENTION: SIGNAL
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESS: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,308
FILING DATE: 09-DEC-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/224,593
FILING DATE: 05-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UTSD:335--1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 418-3000
TELEFAX: 474-7577
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1956 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-762-308-10

Query Match 84.0%; Score 15.8; DB 2; Length 1956;
Best Local Similarity 90.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 99gtcgtcagcagggggg 20
|||||
Db 1446 GGGTCGTGCGAGGGGGG 1427

RESULT 4
US-09-462-284-1/c
Sequence 1, Application US/09462284
Patent No. 6309868
GENERAL INFORMATION:
APPLICANT: Nestec S.A.
APPLICANT: Monod, Michel
APPLICANT: Doumas, Agnes
APPLICANT: Affolter, Michael
APPLICANT: Van Den Broek, Peter
TITLE OF INVENTION: CLONING OF THE
TITLE OF INVENTION: PROLYL-DIPEPTIDYL-PEPTIDASE FROM
FILE REFERENCE: 8265-298
CURRENT APPLICATION NUMBER: US/09/462,284
CURRENT FILING DATE: 2000-01-03
NUMBER OF SEQ ID NOS: 9

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:06:30 ; Search time 277.54 seconds
(without alignments)
17.701 Million cell updates/sec

Title: US-09-672-126-33

Perfect score: 20
Sequence: 1 gggctgcgcagcagggggg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_patents_NA:*
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.4	92.0	400	US-08-881-450A-1	Sequence 1, Appl
C 2	18.4	92.0	5658	US-08-881-450A-23	Sequence 23, Appl
C 3	16.8	84.0	1956	US-08-762-308-10	Sequence 10, Appl
C 4	15.8	79.0	5496	US-09-462-284-1	Sequence 1, Appl
C 5	15.8	79.0	6854	US-09-194-005-7	Sequence 7, Appl
C 6	15.2	76.0	60	US-08-349-696-4	Sequence 4, Appl
C 7	15.2	76.0	60	US-08-233-009-4	Sequence 4, Appl
C 8	15.2	76.0	60	US-08-360-231-4	Sequence 4, Appl
C 9	15.2	76.0	60	US-09-080-704A-4	Sequence 4, Appl
C 10	15.2	76.0	975	US-08-440-846-1	Sequence 1, Appl
C 11	15.2	76.0	1403	US-08-202-044-1	Sequence 1, Appl
C 12	15.2	76.0	1403	US-08-751-344B-1	Sequence 1, Appl
C 13	15.2	76.0	1614	US-08-583-672-1	Sequence 1, Appl
C 14	15.2	76.0	1688	US-08-442-884-2	Sequence 2, Appl
C 15	15.2	76.0	1898	US-08-342-411A-1	Sequence 1, Appl
C 16	15.2	76.0	28958	US-08-258-261B-6	Sequence 6, Appl
C 17	15.2	76.0	28958	US-08-456-837-6	Sequence 6, Appl
C 18	15.2	76.0	28958	US-08-457-342-6	Sequence 6, Appl
C 19	15.2	76.0	28958	US-08-457-646A-6	Sequence 6, Appl
C 20	15.2	76.0	28958	US-08-458-076A-6	Sequence 6, Appl
C 21	15.2	76.0	28958	US-08-764-233A-4	Sequence 4, Appl
C 22	15.2	76.0	28958	US-08-457-335A-6	Sequence 6, Appl
C 23	15.2	76.0	28958	US-08-729-214-6	Sequence 6, Appl
C 24	15.2	76.0	28958	US-09-028-934-6	Sequence 6, Appl
C 25	15.2	76.0	49377	US-08-764-233A-1	Sequence 1, Appl
C 26	15.2	75.0	1801	PCT-US95-02455-1	Sequence 1, Appl
C 27	14.8	74.0	277	US-08-997-080-99	Sequence 99, Appl

C 28	14.8	74.0	277	US-08-997-362-99	Sequence 99, Appl
C 29	14.8	74.0	277	US-08-873-970-99	Sequence 99, Appl
C 30	14.8	74.0	277	US-09-095-855-99	Sequence 99, Appl
C 31	14.8	74.0	277	US-09-324-542-99	Sequence 99, Appl
C 32	14.8	74.0	1716	US-09-321-961-4	Sequence 4, Appl
C 33	14.8	74.0	13987	US-08-804-227C-13	Sequence 13, Appl
C 34	14.8	74.0	43280	US-08-804-227C-1	Sequence 1, Appl
C 35	14.8	74.0	68750	US-09-335-409-1	Sequence 1, Appl
C 36	14.8	74.0	68750	US-09-568-102-1	Sequence 1, Appl
C 37	14.8	74.0	68750	US-09-567-969-1	Sequence 1, Appl
C 38	14.8	74.0	68750	US-09-568-480-1	Sequence 1, Appl
C 39	14.8	74.0	68750	US-09-568-486-1	Sequence 1, Appl
C 40	14.8	74.0	68750	US-09-568-472-1	Sequence 1, Appl
C 41	14.4	72.0	420	US-08-470-179-162	Sequence 162, App
C 42	14.4	72.0	23673	US-09-773-816-1	Sequence 1, Appl
C 43	14.4	72.0	4411529	US-09-103-840A-1	Sequence 1, Appl
C 44	14.2	71.0	150	US-07-969-931-14	Sequence 14, Appl
C 45	14.2	71.0	150	US-07-855-417A-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-08-881-450A-1/C
: Sequence 1, Application US/08881450A
: Patent No. 6274310
: GENERAL INFORMATION:
: APPLICANT: Habener, J.F. and Stoffers, D.A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESS: Banner & Witcoff, Inc.
: STREET: One Financial Center
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Wordperfect 6.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/881,450A
: FILING DATE: June 24, 1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Kathleen M. Williams
: REGISTRATION NUMBER: 34,380
: REFERENCE/DOCKET NUMBER: 11275/7823
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-345-9100
: TELEFAX: 617-345-9111
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 400 nucleotides
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: genomic DNA
: FEATURE:
: NAME/KEY: human IPF-1 gene
: LOCATION: exon 1
: US-08-881-450A-1

Query Match 92.0%; Score 18.4; DB 4; Length 400;
Best local Similarity 95.0%; Pred. No. 6.6;

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes that this is crucial for ensuring transparency and accountability in the organization's operations.

2. The second part of the document outlines the specific procedures and protocols that must be followed when conducting financial transactions. It details the steps for initiating a transaction, the required approvals, and the documentation needed to support each entry.

3. The third part of the document addresses the role of the accounting department in monitoring and reporting on the organization's financial health. It describes how the department tracks expenses, manages budgets, and provides regular updates to the management team.

4. The final part of the document discusses the importance of regular audits and reviews. It explains how these processes help to identify potential issues, ensure compliance with relevant laws and regulations, and provide an opportunity for continuous improvement in the organization's financial management practices.

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:375117"
/clone_lib="Soares mouse embryo NDM13.5 14.5"
/sex="unknown"
/issue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAACTGGAGCGCGCGGAATTTTCTTTTCTTTTCTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2 ]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M. Fatima Bonaldo."
BASE COUNT      111 a      136 c      153 g      94 t
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 494;
Best Local Similarity 90.0%; Pred. No. 9.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gggtcgtcgacgagggggg 20
||||| |||||||
DB 216 gggtcgtcgacgagggggc 197

```

```

RESULT 15
BF118096/C      506 bp      mRNA      linear      EST 29-DEC-2000
LOCUS      uz11g10.y1 NCI-CGAP_Mam5 Mus musculus CDNA clone IMAGE:3668802 5'
DEFINITION      similar to SW:TNRL_MOUSE_P25118 TUMOR NECROSIS FACTOR RECEPTOR 1
PRECURSORS      , mRNA sequence.
ACCESSION      BF118096
VERSION      BF118096.1 GI:10987572
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 506)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/resources.shtml
MG1:1429570
Seq primer: -40RP from Gibco
High quality sequence stop: 439.
Location/Qualifiers
1..506
/organism="Mus musculus"
/strain="C57/B6"
/db_xref="taxon:10090"
/clone="IMAGE:3668802"
/clone_lib="NCI-CGAP_Mam5"
/tissue_type="tumor, gross tissue"

```

```

FEATURES
source

```

```

/dev_stage="7 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: PCWV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT      106 a      157 c      147 g      96 t
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 506;
Best Local Similarity 90.0%; Pred. No. 9.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gggtcgtcgacgagggggg 20
||||| |||||||
DB 404 gggtcgtcgacgagggggc 385

```

Search completed: August 10, 2002, 02:11:26
 Job time: 13147 sec

DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 224A13 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION AL175894

VERSION AL175894.1 GI:7813951

KEYWORDS GSS: genome survey sequence.

SOURCE Tetraodon nigroviridis.

ORGANISM Tetraodon nigroviridis.

REFERENCE 1 (bases 1 to 474)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

TITLE Unpublished

JOURNAL 2 (bases 1 to 474)

REFERENCE Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

TITLE Unpublished

JOURNAL 3 (bases 1 to 474)

REFERENCE Genoscope.

AUTHORS Direct Submission

COMMENT Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

source

1..474

/organism="Tetraodon nigroviridis"

/db_xref="taxon:99883"

/clone="224A13"

/clone_1lb="g"

/note="Genoscope sequence ID : CGAG224A07LPI-end : T7"

BASE COUNT 117 a 132 c 139 g 78 t 8 others

ORIGIN

Query Match 84.0% Score 16.8; DB 12; Length 474;
Best Local Similarity 90.0% Pred. No. 9.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gggtcgtcgcagagggggg 20
||||||| |||||||
Db 164 GGGTCGTCGGGAGGGGGG 145

RESULT 13

AG126381 485 bp DNA linear GSS 04-NOV-2001

LOCUS Pan troglodytes DNA, clone: PTB-136019.F, genomic survey sequence.

DEFINITION AG126381

ACCESSION AG126381.1 GI:16655546

KEYWORDS GSS: GSS (genome survey sequence).

SOURCE Pan troglodytes male lymphoblast DNA, clone_1lb:PTB Chimpanzee Male BAC library clone:PTB-136019.F.

ORGANISM Pan troglodytes

REFERENCE 1 (sites)
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Pan.
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Pan.

AUTHORS Fujiiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of library PTB

TITLE Unpublished

JOURNAL 2 (bases 1 to 485)

REFERENCE Fujiiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,

Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Direct Submission

JOURNAL Submitted (02-AUG-2001) Aseo Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimbesc@riken.go.jp, URL:http://npg.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

COMMENT PRIMERS

Sequencing: -21M13

LIBRARY

Vector : PKS145

R.site 1 : SacI

R.site 2 : SacI

FEATURES

source

1..485

/organism="Pan troglodytes"

/db_xref="taxon:9598"

/clone="PTB-136019.F"

/sex="male"

/cell_type="lymphoblast"

/clone_1lb="PTB Chimpanzee Male BAC Library"

BASE COUNT 137 a 91 c 232 g 13 t 12 others

ORIGIN

Query Match 84.0% Score 16.8; DB 12; Length 485;
Best Local Similarity 90.0% Pred. No. 9.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gggtcgtcgcagagggggg 20
||||| |||||||||
Db 173 GGGCGCCGACGAGGGGGG 192

RESULT 14

W65172/c 494 bp mRNA linear EST 10-JUN-1996

LOCUS m84f11.r1 Soares mouse embryo NbMe13.5 14.5 MUS musculus cDNA

DEFINITION clone IMAGE:375117 5' similar to gb:W65121 TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (HUMAN); gb:M59377 Murine tumor necrosis factor II receptor (MOUSE);, mRNA sequence.

ACCESSION W65172

VERSION W65172.1 GI:1371410

KEYWORDS EST.

SOURCE EST.

ORGANISM house mouse.

REFERENCE Mus musculus

AUTHORS Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi; Mammalia: Eutheria: Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 494)

Geisel,S., Kucaba,T., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
The Washu-HMI Mouse EST Project

TITLE Unpublished (1996)

JOURNAL Contact: Maria M/Mouse EST Project

COMMENT Washu-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: ESTPrimer
High quality sequence stop: 343.

FEATURES

source

1..494

VERSION AL051269.1 GI:4931484
 KEYWORDS GSS.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 976)
 REFERENCE Genoscope.
 AUTHORS Direct Submission
 TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 Web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osoegawa and
 Aaron Mammoser in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPci-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain y2; cn bw sp, the same strain used for the BDGP's
 p1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
 location/Qualifiers
 1..976
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_lib="RPci-98"
 /clone="BACR010106"
 /note="end : 97"
 BASE COUNT 249 a 185 c 220 g 302 t 20 others
 ORIGIN
 Query Match 85.0%; Score 17; DB 12; Length 976;
 Best Local Similarity 100.0%; Pred. No. 8.2e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 tcgtcagcagagggggg 20
 ||||||||||||||||
 Db 318 TCGTCAGCAGGGGGG 302
 RESULT 8
 LOCUS CNS001CL 1201 bp DNA linear GSS 04-JUN-1999
 DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
 BACR03K12 of RPci-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL060206
 AL060206.1 GI:4939299
 GSS.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1201)
 REFERENCE Genoscope.
 AUTHORS Direct Submission
 TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 Web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osoegawa and
 Aaron Mammoser in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPci-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain y2; cn bw sp, the same strain used for the BDGP's
 p1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
 location/Qualifiers
 1..1201
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_lib="RPci-98"
 /clone="BACR03K12"
 /note="end : 97"
 BASE COUNT 269 a 244 c 254 g 371 t 63 others
 ORIGIN
 Query Match 85.0%; Score 17; DB 12; Length 1201;
 Best Local Similarity 100.0%; Pred. No. 8.4e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 tcgtcagcagagggggg 20
 ||||||||||||||||
 Db 309 TCGTCAGCAGGGGGG 293
 RESULT 9
 LOCUS BE494472 324 bp mRNA linear EST 02-AUG-2000
 DEFINITION WHEI256.H01.002ZS Secale cereale anthr cDNA library Secale cereale
 CDNA clone WHEI256.H01.002, mRNA sequence.
 BE494472
 BE494472.1 GI:9661065
 EST.
 VERSION BE494472.1 GI:9661065
 KEYWORDS EST.
 SOURCE rye.
 ORGANISM Secale cereale
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Secale.
 1 (bases 1 to 324)
 REFERENCE Anderson,O.D., Butler,E., Chao,S., Choi,D.W., Close,T.J., Fenton
 R.D., Gustafson,J.P., Han,P.S., Hala,C.C., Kang,Y., Iazo,G.R.,
 Miller,R., Rausch,C.J., Ross,K., Seaton,C.L. and Tong,J.C.
 The structure and function of the expressed portion of the wheat
 genomes - Anthr cDNA library from rye
 Unpublished (2000)
 JOURNAL Contact: Olin Anderson
 COMMENT US department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105595773
 Fax: 5105595818
 Email: oanderson@w.usda.gov
 Sequence have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: Stratagene SK primer.
 Location/Qualifiers
 1..324
 /organism="Secale cereale"
 /cultivar="Blanco"
 /db_xref="taxon:4550"
 /clone="WHEI256.H01.002"
 /clone_lib="Secale cereale anthr cDNA library"
 /tissue_type="Anthr"
 /dev_stage="Adult plant before anthesis"
 /lab_host="E. coli SOLR"
 /note="Vector: lambda Uni-ZAP XR, excised phagemid;
 Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
 greenhouse. Anthr were harvested and pooled from early

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

REFERENCE 1 (bases 1 to 348)

AUTHORS Sasaki, T. and Yamamoto, K.

TITLE Rice cDNA from young root (2000)

JOURNAL Unpublished (2000)

COMMENT Contact: Takuji Sasaki

National Institute of Agrobiological Resources

Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki

305-8602, Japan

Tel: 81-298-38-7441

Fax: 81-298-38-7468

Email: tsasaki@abr.affrc.go.jp, URL: http://jsg.dna.affrc.go.jp/
PROJECT="RGP"

RI0677_22.

FEATURES

source

1. 348

/organism="Oryza sativa"

/strain="Nipponbare"

/db_xref="taxon:4530"

/clone="RI0677"

/clone_lib="Rice cDNA from young root"

/issue_type="young root"

47 a 106 c 146 g 46 t 3 others

BASE COUNT

47 a 106 c 146 g 46 t 3 others

ORIGIN

Query Match

Best Local Similarity 90.0%; Score 18; DB 9; Length 348;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 gtcgtgcagcagggggg 20

|||||

Db 110 gtcgtgcagcagggggg 127

RESULT 5

LOCUS

AV641692 143 bp mRNA linear EST 15-DEC-2000

DEFINITION AV641692 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii

AV641692 cDNA clone HCL038906_r 5', mRNA sequence.

AV641692.1 GI:10785020

VERSION

KEYWORDS

SOURCE

ORGANISM Chlamydomonas reinhardtii.

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 143)

Asanizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,

Nakamura, Y., and Tabata, S. Generation of expressed sequence tags from low-CO2 and high-CO2

adapted cells of Chlamydomonas reinhardtii

20539644

CONTACT: Erika Asanizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asanizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source

1. 143

/organism="Chlamydomonas reinhardtii"

/strain="C9"

/db_xref="taxon:3055"

/clone="HCL038906_r"

/clone_lib="Chlamydomonas reinhardtii 5% CO2"

/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from cells cultured

in a medium with bubbling air containing 5% carbon

BASE COUNT

16 a 58 c 53 g 16 t

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 143;

Best Local Similarity 94.7%; Pred. No. 5e+03;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ggtcgtgcagcagggggg 19

|||||

Db 138 GGTCTGTGGCGAGGGGGG 120

RESULT 6 1114 bp mRNA linear EST 19-OCT-2000

LOCUS BF101822/c 601753172F1 NC1_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3980800 5',

DEFINITION BF101822

RNA sequence.

ACCESSION BF101822

VERSION BF101822.1 GI:10884348

KEYWORDS

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1114)

NIH-MGC http://mgc.nci.nih.gov/.

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cga@bbs-rmail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM9176 row: P column: 17

High quality sequence stop: 656.

Location/Qualifiers

1. 1114

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3980800"

/clone_lib="NCI_CGAP_Mam1"

/issue_type="tumor, biopsy sample"

/dev_stage="10 months, virgin"

/note="Organ: mammary; Vector: PCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

BASE COUNT 354 a 329 c 245 g 186 t

ORIGIN

Query Match

Best Local Similarity 87.0%; Score 17.4; DB 10; Length 1114;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 gtcgtgcagcagggggg 20

|||||

Db 904 GGTCTGTGGCGAGGGGGG 886

RESULT 7

LOCUS CNS004CM/ 976 bp DNA linear GSS 03-JUN-1999

DEFINITION CNS004CM Drosophila melanogaster genome survey sequence T7 end of BAC #

BACR10106 of RPCI-98 library from Drosophila melanogaster (fruit

fly), genomic survey sequence.

ACCESSION AL051269


```

/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/Note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110. E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

BASE COUNT      103 a      256 c      190 g      74 t
ORIGIN

Query Match      92.0%; Score 18.4; DB 10; Length 623;
Best Local Similarity 95.0%; Pred. No. 2.5e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  gggctgcgcagcagggggg 20
          ||||||| |||||||
Db      340 GGGTCGTCGCGCAGGCGGCG 321

RESULT 2
LOCUS      B1715297
DEFINITION B1715297 626 bp mRNA 5' linear EST 19-SEP-2001
            1c31c07.y1 HR85 islet Homo sapiens CDNA 5' similar to SW:IPFL_HUMAN
ACCESSION  P52945 INSULIN PROMOTER FACTOR 1, mRNA sequence.
VERSION     B1715297
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 626)
AUTHORS      Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishke,I., Seearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hiller,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blisstein,A.,
            Schmitt,A., Rheising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvill,I.R., Williams,T.,
            Jackson,Y. and Bowers,Y.
            Endocrine Pancreas Consortium
            Unpublished (2000)
            Other ESTs: 1c31c07.x1
TITLE        Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
JOURNAL      Endocrine Pancreas Consortium
COMMENT      Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biochem.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@imgate.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 481.
FEATURES
    source
        1..626
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="HR85 islet"
            /tissue_type="Purified pancreatic islet"
            /lab_host="DH10B"
            /Note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
            Size-selected on agarose gel. Average insert size ~1kb. 5'
            XhoI site was destroyed after directional cloning.
            Amplified once. Contact information: Hiroshi Inoue, MD,
            Metabolism Div. (Alan Permutt Lab), Washington University
            School of Medicine, Box 8127, 660 South Euclid Ave., St.
            Louis, MO 63110. E-mail: hinoue@imgate.wustl.edu, Tel:
            314-362-1916, Fax: 314-747-2692."

```

```

Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110. E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

BASE COUNT      102 a      255 c      198 g      71 t
ORIGIN

Query Match      92.0%; Score 18.4; DB 10; Length 626;
Best Local Similarity 95.0%; Pred. No. 2.5e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  gggctgcgcagcagggggg 20
          ||||||| |||||||
Db      358 GGGTCGTCGCGCAGGCGGCG 339

RESULT 3
LOCUS      AU078367
DEFINITION AU078367 Rice cDNA from young root Oryza sativa cDNA clone
            R10677.1A, mRNA sequence.
ACCESSION  AU078367
VERSION     AU078367.1 GI:5900712
KEYWORDS    EST.
SOURCE      Oryza sativa.
ORGANISM    Oryza sativa
REFERENCE    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzoae; Oryza.
            1 (bases 1 to 344)
AUTHORS      Yamamoto,K. and Sasaki,T.
TITLE        Rice cDNA from young root
JOURNAL      Unpublished (1999)
COMMENT      Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasak@abrr.affrc.go.jp, URL: http://irg.dna.affrc.go.jp/
            PROJECT "RGP".
FEATURES
    source
        1..344
            /organism="Oryza sativa"
            /strain="Nipponbare"
            /db_xref="taxon:4530"
            /clone_lib="R10677.1A"
            /clone_lib="Rice cDNA from young root"
            /tissue_type="young root"
BASE COUNT      47 a      104 c      146 g      46 t      1 others
ORIGIN

Query Match      90.0%; Score 18; DB 9; Length 344;
Best Local Similarity 100.0%; Pred. No. 3.3e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3  gtcgtgcagcagggggg 20
          ||||||| |||||||
Db      111 GTCGTCGCGCAGGCGGCG 128

RESULT 4
LOCUS      AU102212
DEFINITION AU102212 Rice cDNA from young root Oryza sativa cDNA clone R10677,
            mRNA sequence.
ACCESSION  AU102212
VERSION     AU102212.1 GI:9886319
KEYWORDS    EST.
SOURCE      Oryza sativa.
ORGANISM    Oryza sativa

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:11:23 ; Search time 9068.22 Seconds

(without alignments)
29.768 Million cell updates/sec

Title: US-09-672-126-33

Perfect score: 20

Sequence: 1 gggtcgcgcagcagggg9999 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues 27472414

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	92.0	623	10	BM352781 ig67a09.y
2	18.4	92.0	626	10	BI715297
3	18	90.0	344	9	AU078367
4	18	90.0	348	9	AU102212
5	17.4	87.0	143	9	AV641692
6	17.4	87.0	1114	10	BF101822
7	17	85.0	976	12	CNS004CM
8	17	85.0	1201	12	CNS001CL
9	16.8	84.0	324	10	BE944472
10	16.8	84.0	391	10	BE974442
11	16.8	84.0	468	10	W08614
12	16.8	84.0	474	12	CNS020DP
13	16.8	84.0	485	12	AG126381
14	16.8	84.0	494	10	W65172
15	16.8	84.0	506	10	BF118096
16	16.8	84.0	516	10	BE426494
17	16.8	84.0	529	9	AU101404

C 18	16.8	84.0	537	9	AA265597	AA265597 m246e05.r
C 19	16.8	84.0	537	9	AA466639	AA466639 ve19g08.r
C 20	16.8	84.0	537	9	BB752495	BB752495 BB752495
C 21	16.8	84.0	539	9	AA020177	AA020177 mb51f04.r
C 22	16.8	84.0	542	9	BB764141	BB764141 BB764141
C 23	16.8	84.0	556	10	W99099	W99099 m192c12.r1
C 24	16.8	84.0	563	10	BM235169	BM235169 K0413E05
C 25	16.8	84.0	563	10	BE692231	BE692231 uc26f06.x
C 26	16.8	84.0	574	9	AA238464	AA238464 mx83d05.r
C 27	16.8	84.0	579	10	BC973301	BC973301 602842523
C 28	16.8	84.0	603	9	AA710879	AA710879 vt54c09.r
C 29	16.8	84.0	609	10	BI966654	BI966654 id57e06.x
C 30	16.8	84.0	611	9	A1622583	A1622583 486058G10
C 31	16.8	84.0	619	10	BF462789	BF462789 UI-M-CGDP
C 32	16.8	84.0	619	10	BE334962	BE334962 us91b07.y
C 33	16.8	84.0	642	12	AG140987	AG140987 Pan t10g1
C 34	16.8	84.0	656	10	BF224858	BF224858 uz19g10.x
C 35	16.8	84.0	675	12	AG139625	AG139625 Pan t10g1
C 36	16.8	84.0	697	10	BI412749	BI412749 602989290
C 37	16.8	84.0	776	9	A1267046	A1267046 u110b06.x
C 38	16.8	84.0	789	10	BF138948	BF138948 601783079
C 39	16.8	84.0	842	10	BI155446	BI155446 602903824
C 40	16.8	84.0	853	10	BI525375	BI525375 602924303
C 41	16.8	84.0	928	10	BS985684	BS985684 2051 NICB
C 42	16.8	84.0	1143	10	BM457175	BM457175 AGENCOURT
C 43	16.8	84.0	1181	10	BI196789	BI196789 602755736
C 44	16.8	84.0	1295	10	BS920000	BS920000 602823396
C 45	16.8	84.0	1439	10	BE868507	BE868507 601444520

ALIGNMENTS

RESULT 1
BM352781/c 623 bp mRNA linear EST 07-JAN-2002
LOCUS ig67a09.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:IPFL_HUMAN
DEFINITION P52945 INSULIN PROMOTER FACTOR 1 ; mRNA sequence.
ACCESSION BM352781 GI:18085139
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 623)
Melton,D., Brown,J., Kently,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Maira,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Rilter,E., Ronko,I., Bennett,J., Cardenas
, M., Gibbons,M., McCann,R., Cole,R., Tsagarelshvili,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Human.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: ig67a09.x1
Contract: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center for information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 481.
Location/Qualifiers
1..623
/organism="Homo sapiens"
/db_xref="taxon:9606"

TITLE
JOURNAL
COMMENT

FEATURES
source

THIS PAGE BLANK (USPTO)

CC Sequences AAS59506-AAS59804 represent DNA molecules encoding
 CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
 CC associated DNA sequences are used in the treatment, prevention and
 CC diagnosis of medical conditions caused by P. acnes. The disorders include
 CC SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and
 CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved
 CC in infections of bone, joints and the central nervous system, however it
 CC is particularly involved in the inflammatory lesions associated with acne
 CC vulgaris. A method for detecting the presence or absence of P. acnes in a
 CC patient comprises contacting a sample with a binding agent that binds to
 CC the proteins of the invention and determining the amount of bound protein
 CC in the sample. The polypeptides may be used as antigens in the production
 CC of antibodies specific for P. acnes proteins. These antibodies can be
 CC used to downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
 CC polypeptides shown in AAU51663-AAU51893 and AAU67535.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC XX

SO Sequence 23128 BP; 4349 A; 6746 C; 7113 G; 4908 T; 12 other;

Query Match 82.0%; Score 16.4; DB 23; Length 23128;
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggtcgtcgacgaggg 18
 |||||

Db 19803 GGTCGTGACGAGCGG 19786

RESULT 15
 AAS59539/C
 ID AAS59539 standard; DNA: 29634 BP.
 AC AAS59539;
 DT 13-FEB-2002 (first entry)
 DE Propionibacterium acnes immunogenic protein encoding DNA #34.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant; ds.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PE 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Claim 1; SEQ ID No 34; 1069pp; English.

XX
 CC Sequences AAS59506-AAS59804 represent DNA molecules encoding
 CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
 CC associated DNA sequences are used in the treatment, prevention and
 CC diagnosis of medical conditions caused by P. acnes. The disorders include
 CC SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and
 CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved
 CC in infections of bone, joints and the central nervous system, however it
 CC is particularly involved in the inflammatory lesions associated with acne
 CC vulgaris. A method for detecting the presence or absence of P. acnes in a
 CC patient comprises contacting a sample with a binding agent that binds to
 CC the proteins of the invention and determining the amount of bound protein
 CC in the sample. The polypeptides may be used as antigens in the production
 CC of antibodies specific for P. acnes proteins. These antibodies can be
 CC used to downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
 CC polypeptides shown in AAU47468-AAU47821.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC XX

SO Sequence 29634 BP; 5743 A; 9719 C; 8691 G; 5479 T; 2 other;

Query Match 80.0%; Score 16; DB 23; Length 29634;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggtcgtcgacgaggg 16
 |||||

Db 10686 GGTCGTGACGAGCGG 10671

Search completed: August 10, 2002, 03:21:56
 Job time: 13687 sec

XX (IOWA) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
DR WPI: 2001-273485/28.
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 59; 338pp; English.
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SO Sequence 20 BP; 0 A; 3 C; 13 G; 4 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 2.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggtcgtcgacgagggggg 20
||||| ||| ||| ||| |||
Db 1 gggtcgtcgtcgagggggg 20

RESULT 13

AAf99742
ID AAF99742 standard; DNA; 21 BP.

XX
AC AAF99742;

XX
DT 12-JUN-2001 (first entry)

XX
DE Immunostimulatory nucleic acid #858.

XX
KW Vaccine; cytostatic; vitucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.

XX
OS Synthetic.

XX
PN WO200122972-A2.

XX
PD 05-APR-2001.

XX
PF 25-SEP-2000; 2000WO-US26383.

XX
PR 25-SEP-1999; 99US-0156113.

XX
PR 27-SEP-1999; 99US-0156135.

XX
PR 23-AUG-2000; 2000US-0227436.

XX
PA (IOWA) UNIV IOWA RES FOUND.

XX
PA (COLE-) COLEY PHARM GMBH.

XX
PI Krieg AM, Schetter C, Vollmer J;

XX
DR WPI: 2001-273485/28.

XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 57; 338pp; English.

XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.

SO Sequence 21 BP; 3 A; 2 C; 12 G; 3 T; 1 other;

Query Match 82.0%; Score 16.4; DB 22; Length 21;
Best Local Similarity 85.0%; Pred. No. 3.4e+02;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggtcgtcgacgagggggg 20
||||| ||| ||| ||| |||
Db 2 gggtcatcgatgagggggg 21

RESULT 14

AA559552/C
ID AAS59552 standard; DNA; 23128 BP.

XX
AC AAS59552;

XX
DT 13-FEB-2002 (first entry)

XX
DE Propionibacterium acnes immunogenic protein encoding DNA #47.

XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant; ds.

XX
OS Propionibacterium acnes.

XX
PN WO200181581-A2.

XX
PD 01-NOV-2001.

XX
PF 20-APR-2001; 2001WO-US12865.

XX
PR 21-APR-2000; 2000US-199047P.

XX
PR 02-JUN-2000; 2000US-208841P.

XX
PR 07-JUL-2000; 2000US-216747P.

XX
PA (CORI-) CORIXA CORP.

XX
PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX
DR WPI: 2001-616774/71.

XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -

XX
PS Claim 1; SEQ ID NO 47; 1069pp; English.

KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytosstatic; nootopic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antineumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antineumatic; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 PS Claim 1; SEQ ID NO 1042; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 CC
 XX
 SQ Sequence 9117 BP; 2130 A; 465 C; 2553 G; 3969 T; 0 other;
 XX
 Query Match 92.0%; Score 18.4; DB 24; Length 9117;
 Best Local Similarity 95.0%; Pred. No. 31;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 gggtcgtcgacgagggggg 20
 ||||||||| |||||||||
 Db 3889 gggtcgtcgacgagggggg 3908
 XX
 RESULT 11
 ID AAF98879 standard; DNA: 20 BP;
 XX
 AC AAF98879;
 XX
 DT 11-JUN-2001 (first entry)
 XX
 DE Immunostimulatory nucleic acid assay control oligo SEQ ID NO: 160.
 XX
 KW Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
 KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT modified_base 1..2
 FT /*tag= a
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"

FT modified_base 15..19
 FT /*tag= b
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 XX
 PN WO200122990-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 27-SEP-2000; 2000WO-US26527.
 XX
 PR 27-SEP-1999; 99US-0156147.
 XX
 PA (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX
 PI Hartmann G, Bratzler RL, Krieg A;
 XX
 DR WPI; 2001-290487/30.
 XX
 PT Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 XX
 PS Example 17; Page 166; 168pp; English.
 XX
 CC The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 CC
 XX
 SQ Sequence 20 BP; 0 A; 3 C; 13 G; 4 T; 0 other;
 XX
 Query Match 84.0%; Score 16.8; DB 22; Length 20;
 Best Local Similarity 90.0%; Pred. No. 2.3e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 gggtcgtcgacgagggggg 20
 ||||||||| |||||||||
 Db 1 gggtcgtcgctgtg99ggg 20
 XX
 RESULT 12
 ID AAF9868 standard; DNA: 20 BP;
 XX
 AC AAF9868;
 XX
 DT 12-JUN-2001 (first entry)
 XX
 DE Immunostimulatory nucleic acid #984.
 XX
 KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX
 OS Synthetic.
 XX
 PD 05-APR-2001.
 XX
 PF 25-SEP-2000; 2000WO-US26383.
 XX
 PR 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.

```

XX 06-APR-1999 (first entry)
DT IPF1 gene exon encoding for ORF1, ORF2 and ORF3.
XX
XX Mature onset diabetes of the young; MODY; insulin promoter factor 1;
XX IPF1; mutation; MODY4; pancreatic disorder; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX CDS 20..391
XX FT /tag= a
XX FT /product= "ORF1 product"
XX FT /note= "translated protein sequence in AAW95596"
XX CDS 70..204
XX FT /tag= b
XX FT /product= "ORF3 product"
XX FT /note= "translated protein sequence in AAW95598"
XX CDS 207..383
XX FT /tag= c
XX FT /product= "ORF2 product"
XX FT /note= "translated protein sequence in AAW95597"
XX
XX MO9859078-A1.
XX
XX 30-DEC-1998.
XX
XX 24-JUN-1998; 98WO-US13467.
XX
XX 24-JUN-1997; 97US-0881450.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX Habener JF, Stoffers DA;
XX
XX WPI; 1999-105636/09.
XX
XX P-PSDB; AAW95596, AAW95597, AAW95598.
XX
XX
XX Detecting heterozygosity for insulin promoter factor 1 - useful to
XX detect the presence of, or predisposition for, mature onset diabetes
XX of the young
XX
XX Disclosure: Page 20; 46pp; English.
XX
XX The invention relates to a new method to screen for mature onset diabetes
XX of the young (MODY). The method comprises detecting a mutation in the
XX gene encoding insulin promoter factor 1 (IPF1), wherein heterozygosity
XX for the mutation is indicative of MODY. The method may be used to
XX determine if a patient with MODY symptoms has MODY4, to assess patients'
XX risk of developing MODY4, to assess the risk of a couple's progeny of
XX inheriting MODY, and to assist in determining the genetic basis for other
XX pancreatic disorders that might result from IPF-1 deficiency. The present
XX sequence represents the exon 1 of IPF1 gene encoding for ORF1, ORF2 and
XX ORF3. The mutation is preferably a C-terminal deletion of the ORF1
XX product at Pro63.
XX
XX Sequence 400 BP; 58 A; 176 C; 118 G; 48 T; 0 other;
XX
XX Query Match 92.0%; Score 18.4; DB 20; Length 400;
XX Best Local Similarity 95.0%; Pred. No. 39;
XX Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 gggctgcgcagcagggggg 20
XX ||||||| |||||||
XX DB 249 gggctgcgcagcagggggg 230
XX
XX RESULT 9
XX AAX01053/c
XX ID AAX01053 standard; DNA; 5658 BP.
XX

```

```

AC AAX01053;
XX
XX 06-APR-1999 (first entry)
DT Nucleotide sequence of IPF1 gene contig 2.
XX
XX Mature onset diabetes of the young; MODY; insulin promoter factor 1;
XX IPF1; mutation; MODY4; pancreatic disorder; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX CDS 1906..2315
XX FT /tag= a
XX FT /product= "ORF-1 product"
XX FT /note= "first coding region"
XX Intron 3316..5658
XX FT /tag= b
XX FT /number= 1
XX
XX MO9859078-A1.
XX
XX 30-DEC-1998.
XX
XX 24-JUN-1998; 98WO-US13467.
XX
XX 24-JUN-1997; 97US-0881450.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX Habener JF, Stoffers DA;
XX
XX WPI; 1999-105636/09.
XX
XX P-PSDB; AAW95596.
XX
XX
XX Detecting heterozygosity for insulin promoter factor 1 - useful to
XX detect the presence of, or predisposition for, mature onset diabetes
XX of the young
XX
XX Disclosure: Pages 26-28; 46pp; English.
XX
XX The invention relates to a new method to screen for mature onset diabetes
XX of the young (MODY). The method comprises detecting a mutation in the
XX gene encoding insulin promoter factor 1 (IPF1), wherein heterozygosity
XX for the mutation is indicative of MODY. The method may be used to
XX determine if a patient with MODY symptoms has MODY4, to assess patients'
XX risk of developing MODY4, to assess the risk of a couple's progeny of
XX inheriting MODY, and to assist in determining the genetic basis for other
XX pancreatic disorders that might result from IPF-1 deficiency. The present
XX sequence represents the genomic DNA sequence of IPF1 gene contig 2.
XX
XX Sequence 5658 BP; 1299 A; 1633 C; 1444 G; 1197 T; 85 other;
XX
XX Query Match 92.0%; Score 18.4; DB 20; Length 5658;
XX Best Local Similarity 95.0%; Pred. No. 32;
XX Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 gggctgcgcagcagggggg 20
XX ||||||| |||||||
XX DB 2135 gggctgcgcagcagggggg 2116
XX
XX RESULT 10
XX ABL33069
XX ID ABL33069 standard; DNA; 9117 BP.
XX
XX ABL33069;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 1042.
XX

```


Query Match 95.0%; Score 19; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 1 gggtcgtcgacgagggg 19
 ||||||||||||||||
 Db 2 gggtcgtcgacgagggg 20

RESULT 6

AAFG9768
 ID AAF99768 standard; DNA: 20 BP.

AC AAF99768;

DT 12-JUN-2001 (first entry)

DE Immunostimulatory nucleic acid #884.

KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;

KM immunostimulatory; tumour; viral infection; bacterial infection;

KW fungal infection; parasitic infection; cancer; asthma;

KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.

OS Synthetic.

PN WO200122972-A2.

PD 05-APR-2001.

PF 25-SEP-2000; 2000WO-US26383.

PR 25-SEP-1999; 99US-0156113.

PR 27-SEP-1999; 99US-0156135.

PR 23-AUG-2000; 2000US-0227436.

PA (IOWA) UNIV IOWA RES FOUND.

PA (COLE-) COLEY PHARM GMBH.

PI Kriegl AM, Schetter C, Vollmer J;

XX WPI; 2001-273485/28.

DR VPI; 2001-273485/28.

XX Vaccinating against tumors, infectious diseases; allergies and asthma

PT using immunostimulatory Py-rich and TG nucleic acids -

PS Claim 101; Page 57; 338pp; English.

XX The present invention relates to a method for stimulating an immune

CC response. The method comprises administering an immunostimulatory nucleic

CC acid to a non-rodent subject in sufficient quantity to stimulate an

CC immune response. The present sequence is one such immunostimulatory

CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich

CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects

CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae

CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,

CC haemophilus, campylobacter, clostridium, Escherichia coli and/or

CC staphylococcus), fungal antigens and/or parasitic antigens. The method is

CC also useful for preventing cancer, asthma, infectious disease, allergy or

CC immune deficiency. The present sequence can also be used to redirect a

CC Th2 to a Th1 immune response and to activate immune cells.

CC Note: the present sequence may have a phosphorothioate backbone.

CC Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 95.0%; Score 19; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 1 gggtcgtcgacgagggg 19
 ||||||||||||||||
 Db 2 gggtcgtcgacgagggg 20

RESULT 7

AAFG9830
 ID AAF99830 standard; DNA: 20 BP.

AC AAF99830;

DT 12-JUN-2001 (first entry)

DE Immunostimulatory nucleic acid #946.

KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;

KM immunostimulatory; tumour; viral infection; bacterial infection;

KW fungal infection; parasitic infection; cancer; asthma;

KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.

OS Synthetic.

PN WO200122972-A2.

PD 05-APR-2001.

PF 25-SEP-2000; 2000WO-US26383.

PR 25-SEP-1999; 99US-0156113.

PR 27-SEP-1999; 99US-0156135.

PR 23-AUG-2000; 2000US-0227436.

PA (IOWA) UNIV IOWA RES FOUND.

PA (COLE-) COLEY PHARM GMBH.

PI Kriegl AM, Schetter C, Vollmer J;

XX WPI; 2001-273485/28.

DR VPI; 2001-273485/28.

XX Vaccinating against tumors, infectious diseases; allergies and asthma

PT using immunostimulatory Py-rich and TG nucleic acids -

PS Claim 101; Page 58; 338pp; English.

XX The present invention relates to a method for stimulating an immune

CC response. The method comprises administering an immunostimulatory nucleic

CC acid to a non-rodent subject in sufficient quantity to stimulate an

CC immune response. The present sequence is one such immunostimulatory

CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich

CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects

CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae

CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,

CC haemophilus, campylobacter, clostridium, Escherichia coli and/or

CC staphylococcus), fungal antigens and/or parasitic antigens. The method is

CC also useful for preventing cancer, asthma, infectious disease, allergy or

CC immune deficiency. The present sequence can also be used to redirect a

CC Th2 to a Th1 immune response and to activate immune cells.

CC Note: the present sequence may have a phosphorothioate backbone.

CC Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 95.0%; Score 19; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 1 gggtcgtcgacgagggg 19
 ||||||||||||||||
 Db 2 gggtcgtcgacgagggg 20

RESULT 8

AAAX1055/C
 ID AAXX1055 standard; DNA: 400 BP.

AC AAXX1055;

XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX
XX Sequence 19 BP; 2 A; 3 C; 12 G; 2 T; 0 other;

Query Match	95.0%	Score 19;	DB 22;	Length 19;
Best Local Similarity	100.0%	Pred. No. 27;		
Matches 19;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	2	ggtcgtcgacgaaggagg	20	
db	1	ggtcgtcgacgaaggagg	19	.

RESULT 4
 AAF99867 AAF99867 standard; DNA; 19 BP.
 AC AC
 AD AAF99867;
 AE 12-JUN-2001 (first entry)
 AF XX
 AG Immunostimulatory nucleic acid #983.
 AH XX
 AI Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 AJ immunostimulatory; tumour; viral infection; bacterial infection;
 AK fungal infection; parasitic infection; cancer; asthma;
 AL infectious disease; allergy; immune deficiency; phosphorochioate; ss.
 AM XX
 AN Synthetic.
 AO WO200122972-A2.
 AP XX
 AQ PN
 AR PD
 AS 05-APR-2001.
 AT XX
 AU 25-SEP-2000; 2000WO-US26383.
 AV XX
 AW 25-SEP-1999; 99US-0156113.
 AX 27-SEP-1999; 99US-0156135.
 AY PR
 AZ 23-AUG-2000; 2000US-0227436.
 BA XX
 BB (IOWA) UNIV IOWA RES FOUND.
 BC PA
 BD (COLE-) COLEY PHARM GMBH.
 BE XX
 BF Kriegl AM, Schetter C, Vollmer J;
 BG WPI; 2001-273485/28.
 BH XX
 BI Vaccinating against tumors, infectious diseases, allergies and asthma
 BJ using immunostimulatory Py-rich and TG nucleic acids -
 BK XX
 BL Claim 101; Page 59; 338pp; English.
 BM XX
 BN The present invention relates to a method for stimulating an immune
 BO response. The method comprises administering an immunostimulatory nucleic
 BP acid to a non-rodent subject in sufficient quantity to stimulate an
 BQ immune response. The present sequence is one such immunostimulatory
 BA nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 BB (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 BC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 BD and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 BE haemophilus, campylobacter, clostridium, Escherichia coli and/or
 BF staphylococcus), fungal antigens and/or parasitic antigens. The method is
 BG also useful for preventing cancer, asthma, infectious disease, allergy or
 BH immune deficiency. The present sequence can also be used to redirect a

CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone
xx
SQ Sequence 19 BP; 2 A; 3 C; 12 G; 2 T; 0 other;

Query Match	95.08;	Score 19;	DB 22;	Length 19;
Best Local Similarity	100.08;	Pred. No. 27;		
Matches	.19;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	2	ggtcgtcagcagggggg	20
Db	1	ggtcgtcagcagggggg	19

RESULT	5
AAF98748	
ID	AAF98748 standard; DNA: 20 BP.

DT 11-JUN-2001 (first entry)
XX
DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 18.
XX
KW Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha.
XX
RN viral infection; phosphorothioate backbone; palindrome; cancer; ds-
XX

XX	Key	Location/Qualifiers
FT	modified_base	1..2
FT		/*tag= a
FT		/mod_base= "OTHER"
FT		/note= "phosphorothioate linkage"
FT	modified_base	15..19
FT		/*tag= b
FT		/mod_base= "OTHER"
FT		/note= "phosphorothioate linkage"
XX		
XX	WO200122990-A2.	
XX		
XX	05-APR-2001.	
XX		
XX	27-SEP-2000; 2000WO-US26527.	
XX		
XX	27-SEP-1999; 99US-0156147.	
XX		
XX	(COLE-) COLEY PHARM GROUP INC.	
PA	(IOWA) UNIV IOWA RES FOUND.	
PI		
PI	Hartmann G, Bratzler RL, Kriegel A;	
XX		
DR	WPI; 2001-290487/30.	
PT		
PT	Improving the efficacy of treatments involving the administration of	
PT	interferon-alpha by co-administering an isolated immunostimulatory	
PT	nucleic acid -	
XX		
PS	Claim 201; Page 103; 168pp; English.	
XX		
CC	The present invention describes an improvement to a method requiring the	
CC	administration of interferon alpha (IFN-alpha), involving administering	
CC	an immunostimulatory nucleic acid (ISNA). The sequences of a number of	
CC	such nucleic acids are also provided. These may comprise oligonucleotides	
CC	with phosphorothioate backbones, palindromes, or G-rich sequences. The	
CC	sequences of the invention are useful in the treatment of proliferative	
CC	diseases, such as cancers, and viral infections. The present sequence is	
CC	an example of an immunostimulatory oligonucleotide.	
XX		
XX		
XX	Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;	

XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.
XX
PI Hartmann G, Bratzler RL, Krieg A;
XX WPI; 2001-290487/30.
DR
XX
PT Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
PS Claim 201; Page 103; 168bp; English.
XX
CC The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX
SQ Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gggtcgtcgcagcagggggg 20
|||||
Db 1 gggtcgtcgcagcagggggg 20

RESULT 2
AAF9866 standard; DNA; 20 BP.
XX
AC AAF9866;
XX
DT 12-JUN-2001 (first entry)
XX
DE Immunostimulatory nucleic acid #982.
XX
KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS Synthetic.
XX
PN WO200122972-A2.
XX
PD 05-APR-2001.
XX
PF 25-SEP-2000; 2000WO-US26383.
XX
PR 25-SEP-1999; 99US-0156113.
PR 27-SEP-1999; 99US-0156135.
PR 23-AUG-2000; 2000US-0227436.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX WPI; 2001-273485/28.
DR
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 59; 338bp; English.

XX The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
XX
SQ Note: the present sequence may have a phosphorothioate backbone.
SQ Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gggtcgtcgcagcagggggg 20
|||||
Db 1 gggtcgtcgcagcagggggg 20

RESULT 3
AAF98764 standard; DNA; 19 BP.
XX
AC AAF98764;
XX
DT 11-JUN-2001 (first entry)
XX
DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 34.
XX
KW Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH modified_base 1..2
FT /*tag= a /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
FT modified_base 14..18
FT /*tag= b /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
XX
PN WO200122990-A2.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26527.
XX
PR 27-SEP-1999; 99US-0156147.
XX
PA (COLE-) COLEY PHARM GROUP INC.
PA (IOWA) UNIV IOWA RES FOUND.
XX
PI Hartmann G, Bratzler RL, Krieg A;
XX WPI; 2001-290487/30.
DR
XX
PT Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
PS Claim 201; Page 103; 168bp; English.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:21:54 ; Search time 1145.36 seconds
(without alignments)
29.980 Million cell updates/sec

Title: US-09-672-126-33

Perfect score: 20
Sequence: 1 gggctgcgcgcagcagggggg 20

Scoring table: IDENTITY NUC
Gapop 10.0 ; Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_032802:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	22	AAF98763 Human IFN-alpha im
2	20	100.0	20	22	AAF98766 Immunostimulatory
3	19	95.0	19	22	AAF98764 Human IFN-alpha im
4	19	95.0	19	22	AAF98767 Immunostimulatory
5	19	95.0	20	22	AAF98748 Human IFN-alpha im
6	19	95.0	20	22	AAF98768 Immunostimulatory
7	19	95.0	20	22	AAF98830 Immunostimulatory
8	18.4	92.0	400	20	AAK01055 IPII gene exon enc
9	18.4	92.0	5658	20	AAK01053 Nucleotide sequenc

10	18.4	92.0	9117	24	ABL33069 Human immune syste
11	16.8	84.0	20	22	AAF98879 Immunostimulatory
12	16.8	84.0	20	22	AAF98868 Immunostimulatory
13	16.4	82.0	21	22	AAF99742 Immunostimulatory
14	16.4	82.0	23	23	AA559552 Propionibacterium
15	16.4	80.0	23	23	AA559539 Propionibacterium
16	15.8	79.0	19	22	AAF98757 Human IFN-alpha im
17	15.8	79.0	19	22	AAF98771 Human IFN-alpha im
18	15.8	79.0	19	22	AAF98840 Immunostimulatory
19	15.8	79.0	20	22	AAF98735 Human IFN-alpha im
20	15.8	79.0	20	22	AAF98736 Human IFN-alpha im
21	15.8	79.0	20	22	AAF98655 Immunostimulatory
22	15.8	79.0	20	22	AAF98871 Immunostimulatory
23	15.8	79.0	20	22	AAF99704 Immunostimulatory
24	15.8	79.0	20	22	AAF99767 Immunostimulatory
25	15.8	79.0	21	22	AAF98747 Human IFN-alpha im
26	15.8	79.0	21	22	AAF99797 Immunostimulatory
27	15.8	79.0	40	21	AA295996 Polynucleotide seq
28	15.8	79.0	538	21	AAF09234 Fusarium venenatum
29	15.8	79.0	1264	21	AAF07554 Aspergillus oryzae
30	15.8	79.0	2313	20	AAK00013 Aspergillus oryzae
31	15.8	79.0	2588	22	AAH19171 Human secreted pro
32	15.8	79.0	2594	22	AAH19200 Human secreted pro
33	15.8	79.0	5496	20	AAK07327 Aspergillus oryzae
34	15.8	79.0	6114	24	ABL32760 Human immune syste
35	15.8	79.0	6186	23	ABL15588 Drosophila melanog
36	15.8	79.0	6854	19	AAK76903 S. glaucescens pst
37	15.8	79.0	10997	23	ABL02488 Drosophila melanog
38	15.8	79.0	27541	22	AAK17185 Streptomyces nours
39	15.8	79.0	125401	22	AAK17186 Streptomyces nours
40	15.4	77.0	717	22	AAK09826 Corn SPFL-related
41	15.4	77.0	1257	22	AAK08893 Mycobacterium kans
42	15.4	77.0	1263	22	AAK08892 Mycobacterium para
43	15.4	77.0	1263	22	AAK08892 Mycobacterium para
44	15.2	76.0	60	14	AAK05406 Antisense primer f
45	15.2	76.0	60	16	AAK29936 Dog AI antisense o

ALIGNMENTS

RESULT 1	
AAF98763	
ID	AAF98763 standard; DNA; 20 BP.
AC	AAF98763;
XX	
DT	11-JUN-2001 (first entry)
XX	
DE	Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 33.
XX	
KW	Immunostimulatory nucleic acid; ISNA: human; interferon alpha; IFN-alpha;
KW	viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX	
OS	Synthetic.
XX	
FT	Key
FT	modified_base
FT	Location/Qualifiers
FT	1..2
FT	/*tag= a
FT	/mod_base= "OTHER"
FT	/note= "phosphorothioate linkage"
FT	14..19
FT	/*tag= b
FT	/mod_base= "OTHER"
FT	/note= "phosphorothioate linkage"
XX	
XX	WO200122990-A2.
PN	
XX	
XX	05-APR-2001.
PD	
XX	
PF	27-SEP-2000; 2000WO-US26527.
XX	
XX	27-SEP-1999; 99US-0156147.

TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 1042 03-JAN-2002;
Epigenomics AG (DE)
FEATURES Location/Qualifiers
source 1..9117
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 2130 a 465 c 2553 g 3969 t
ORIGIN

Query Match 92.08; Score 18.4; DB 6; Length 9117;
Best Local Similarity 95.08; Pred. No. 6.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 gggtcgtcgacgagggggg 20
||||||| |||||||
Db 3889 GGTCGTCGGGAGGGGGG 3908

Search completed: August 10, 2002, 02:58:46
Job time: 15732 sec


```

FEATURES
  source
    Location/Qualifiers
      1..697
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
BASE COUNT 100 a 282 c 225 g 90 t
ORIGIN
  1 gggtcgtcgcagcagggggg 20
  |||||
  512 gggtcgtcgcagcagggggg 493

Query Match 92.0%; Score 18.4; DB 9; Length 697;
Best Local Similarity 95.0%; Pred. NO. 1.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggtcgtcgcagcagggggg 20
  |||||
  512 gggtcgtcgcagcagggggg 493

Db 512 gggtcgtcgcagcagggggg 493

RESULT 10
HSU35632/c
LOCUS
DEFINITION Human insulin promoter factor 1 (PDX-1) mRNA, complete cds.
ACCESSION U35632
VERSION U35632.1 GI:1197837
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
  1 (bases 1 to 942)
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Stoffel, M., Stein, R., Wright, C.V., Espinosa, R. III, Le Beau, M.M., and Bell, G.I.
TITLE Localization of human homeodomain transcription factor insulin
  promoter factor 1 (IPF1) to chromosome band 13q12.1
JOURNAL Genomics 28 (1), 125-126 (1995)
MEDLINE 96070447
REFERENCE
  2 (bases 1 to 942)
  Wright, C.V.E.
  Direct Submission
  Submitted (07-SEP-1995) Christopher V. E. Wright, Cell Biology,
  Vanderbilt University Medical School, 1161 21st Ave S, Nashville,
  TN 37232-2175, USA
  On Feb 21, 1996 this sequence version replaced gi:1017737.
FEATURES
  source
    Location/Qualifiers
      1..942
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="13"
      /map="13q12.1"
      1..942
      /gene="PDX-1"
      /gene="PDX-1"
      15..866
      /standard_name="pancreas/duodenum homeobox gene-1"
      /note="IPF-1; homeodomain protein; similar to human STF-1;
      PIR Accession Number A47750, IDX-1 in Rattus norvegicus;
      PIR Accession Number S42634, IPF-1 in Mus musculus; PIR
      Accession Number S39581, and X1Hbox8 in Xenopus laevis
      encoded by Genbank Accession Number X16849"
      /codon_start=1
      /product="insulin promoter factor 1"
      /protein_id="AAA88820.1"
      /db_xref="GI:1197838"
      /translation="MNGEEQYAATQLYKDCAFORGAPEFSAPPACLYMGROPPP
      PPFPPGALGALGEGSPDISPEVPLADDPAYAHLLHPLPAQLALPHPPAGFPPE
      GAEPGLVEPNRVLQPFPMKSTKAHWKGWAGGAYAAEPEENKRTTAYTRAOLLE
      LEKEFLNRYISRRPRVELAVMLNTERHIKIWFONRRMKWKEDKRGGTAVGGG
      GVAREQDCAVTSGSELLALPPPPPGAVPPAAPVAAREGLPGLSAPQSSVAP
      RRPQEP"
BASE COUNT 159 a 349 c 320 g 114 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 942;
Best Local Similarity 95.0%; Pred. NO. 1.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggtcgtcgcagcagggggg 20
  |||||
  512 gggtcgtcgcagcagggggg 493

Db 512 gggtcgtcgcagcagggggg 493

RESULT 11
HUMPDX1/c
LOCUS
DEFINITION Homo sapiens insulin promoter factor 1 (PDX1) gene, exon 1.
ACCESSION AF035259
VERSION AF035259.1 GI:2665698
KEYWORDS
SEGMENT
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
  1 (bases 1 to 992)
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Hara, M., Lindner, T.H., Paz, V.P., Wang, X., Iwasaki, N. and Bell, G.I.
TITLE Mutations in the pancreatic and duodenal homeobox gene (PDX1) are
  not a common cause of MODY/Early-onset Type 2 diabetes in Japanese
  Unpublished
JOURNAL
REFERENCE
  2 (bases 1 to 992)
  Hara, M., Lindner, T.H., Paz, V.P., Wang, X., Iwasaki, N. and Bell, G.I.
  Direct Submission
  Submitted (19-NOV-1997) Biochemistry, Molecular Biology, and
  Medicine, The University of Chicago, Howard Hughes Medical
  Institute, 5841 South Maryland Avenue, MC1028, Chicago, IL 60637,
  USA
FEATURES
  source
    Location/Qualifiers
      1..992
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="13"
      /map="13q12.1"
      /clone_lib="PAC library, Genome Systems Inc."
      <283..688
      /gene="PDX1"
      /number=1
      689..>992
      /gene="PDX1"
      /number=1
BASE COUNT 144 a 399 c 288 g 161 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 992;
Best Local Similarity 95.0%; Pred. NO. 1.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggtcgtcgcagcagggggg 20
  |||||
  512 gggtcgtcgcagcagggggg 493

Db 512 gggtcgtcgcagcagggggg 493

RESULT 12
HSU30329/c
LOCUS
DEFINITION Human insulin promoter factor 1 (IPF1) mRNA, complete cds.
ACCESSION U30329
VERSION U30329.1 GI:929922
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
  1 (bases 1 to 1428)
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Inoue, H., Tanizawa, Y. and Permutt, M.A.
TITLE Isolation, characterization, and chromosomal mapping of the human
  insulin promoter factor 1 (IPF1) gene

```

```
source
1. .20
/organism="synthetic construct"
/db_xref="taxon:32630" 2 t
BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gggctgctgcagcagggggg 19
|||||
Db 2 GGGTCGTCGACGAGGGGGG 20

RESULT 6
AX104844
LOCUS AX104844 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1036 from Patent WO0122972.
ACCESSION AX104844
VERSION AX104844.1 GI:13921041
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence...
REFERENCE
1 (bases 1 to 20)
Krieg,A.M., Schetter,C. and Vollmer,J.C.
AUTHORS Immunostimulatory nucleic acids
TITLE Patent: WO 0122972-A 1036 05-APR-2001;
JOURNAL UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES
Location/Qualifiers
source
1. .20
/organism="synthetic construct"
/db_xref="taxon:32630" 2 t
BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gggctgctgcagcagggggg 19
|||||
Db 2 GGGTCGTCGACGAGGGGGG 20

RESULT 7
AX105120
LOCUS AX105120 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 18 from Patent WO0122990.
ACCESSION AX105120
VERSION AX105120.1 GI:13921270
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 20)
Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
AUTHORS Methods related to immunostimulatory nucleic acid-induced
TITLE interferon
JOURNAL Patent: WO 0122990-A 18 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
Location/Qualifiers
source
1. .20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
misc_feature 1. .2
/note="Backbone has phosphorothioate linkages."
```

```
misc_feature 3. .14
/note="Backbone has phosphodiester linkages."
misc_feature 15. .19
/note="Backbone has phosphorothioate linkages."
misc_feature 20
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gggctgctgcagcagggggg 19
|||||
Db 2 GGGTCGTCGACGAGGGGGG 20

RESULT 8
AR164570/c
LOCUS AR164570 400 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6274310.
ACCESSION AR164570
VERSION AR164570.1 GI:16237639
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
1 (bases 1 to 400)
Habener,J.F. and Stoffers,D.A.
AUTHORS Compositions and methods for detecting pancreatic disease
TITLE Patent: US 6274310-A 1 14-AUG-2001;
JOURNAL Location/Qualifiers
FEATURES
source
1. .400
/organism="unknown"
BASE COUNT 58 a 176 c 118 g 48 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 400;
Best Local Similarity 95.0%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gggctgctgcagcagggggg 20
|||||
Db 249 GGGTCGTCGCGAGGGGGG 230

RESULT 9
S82168S1/c
LOCUS S82168S1 697 bp DNA linear PRI 12-FEB-1997
DEFINITION IPF-1-insulin promoter factor 1 [human, Genomic, 697 nt, segment 1
of 2].
ACCESSION S82168
VERSION S82168.1 GI:1839455
KEYWORDS
SEGMENT i of 2
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 697)
Inoue,H., Riggs,A.C., Tanizawa,Y., Ueda,K., Kuwano,A., Liu,L.,
AUTHORS Donis-Keller,H. and Permutt,M.A.
TITLE Isolation, characterization, and chromosomal mapping of the human
insulin promoter factor 1 (IPF-1) gene
JOURNAL Diabetes 45 (6), 789-794 (1996)
MEDLINE 96220081
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 177999] from the original journal article.
This sequence comes from Fig. 1.
Map location: 13q12(12.1).
```



```
Query Match      100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggctcgcacgagggggg 20
   |||||
Db 1 GGGTCGTCGACGAGGGGGG 20

RESULT 2
AX105135
LOCUS AX105135 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 33 from Patent WO0122990.
ACCESSION AX105135
VERSION AX105135.1 GI:13921285
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
        interferon
JOURNAL Patent: WO 0122990-A 33 05-APR-2001;
        Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
        FOUNDATION (US)
FEATURES
    source      Location/Qualifiers
    1..20
    /organism="synthetic construct"
    /db_xref="taxon:32630"
    /note="Synthetic Oligonucleotide"
    misc_feature 1..2
    /note="Backbone has phosphorothioate linkages."
    misc_feature 3..13
    /note="Backbone has phosphodiester linkages."
    misc_feature 14..19
    /note="Backbone has phosphorothioate linkages."
    misc_feature 20
    /note="Backbone has phosphodiester linkages."
BASE COUNT      2 a 3 c 13 g 2 t
ORIGIN
1 gggctcgcacgagggggg 20
|||
1 GGGTCGTCGACGAGGGGGG 20

Query Match      100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gggctcgcacgagggggg 20
   |||||
Db 1 GGGTCGTCGACGAGGGGGG 20

RESULT 3
AX104881
LOCUS AX104881 19 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1073 from Patent WO0122972.
ACCESSION AX104881
VERSION AX104881.1 GI:13921078
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 19)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 1073 05-APR-2001;
        UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
        GmbH (DE)
FEATURES
    source      Location/Qualifiers
    1..19
    /organism="synthetic construct"
    /db_xref="taxon:32630"
```

```
BASE COUNT      2 a 3 c 12 g 2 t
ORIGIN
1 ggtcgtcgcacgagggggg 20
|||
1 GGTCTCGTCGACGAGGGGGG 19

Query Match      95.0%; Score 19; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggtcgtcgcacgagggggg 20
   |||||
Db 1 GGTCTCGTCGACGAGGGGGG 19

RESULT 4
AX105136
LOCUS AX105136 19 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 34 from Patent WO0122990.
ACCESSION AX105136
VERSION AX105136.1 GI:13921286
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 19)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
        interferon
JOURNAL Patent: WO 0122990-A 34 05-APR-2001;
        Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
        FOUNDATION (US)
FEATURES
    source      Location/Qualifiers
    1..19
    /organism="synthetic construct"
    /db_xref="taxon:32630"
    /note="Synthetic Oligonucleotide"
    misc_feature 1..2
    /note="Backbone has phosphorothioate linkages."
    misc_feature 3..13
    /note="Backbone has phosphodiester linkages."
    misc_feature 14..18
    /note="Backbone has phosphorothioate linkages."
    misc_feature 19
    /note="Backbone has phosphodiester linkages."
BASE COUNT      2 a 3 c 12 g 2 t
ORIGIN
1 ggtcgtcgcacgagggggg 20
|||
1 GGTCTCGTCGACGAGGGGGG 19

Query Match      95.0%; Score 19; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggtcgtcgcacgagggggg 20
   |||||
Db 1 GGTCTCGTCGACGAGGGGGG 19

RESULT 5
AX104781
LOCUS AX104781 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 973 from Patent WO0122972.
ACCESSION AX104781
VERSION AX104781.1 GI:13920978
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 973 05-APR-2001;
        UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
        GmbH (DE)
FEATURES
    source      Location/Qualifiers
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:58:44 ; Search time 2778.35 Seconds
(without alignments)
150.640 Million cell updates/sec

Title: US-09-672-126-33

Perfect score: 20

Sequence: 1 gggctgcgcagcagggggg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vl.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vl.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htgo.inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	20	100.0	20	6	AX104880	Sequence
2	20	100.0	20	6	AX105135	Sequence
3	19	95.0	19	6	AX104881	Sequence
4	19	95.0	19	6	AX105136	Sequence
5	19	95.0	20	6	AX104781	Sequence
6	19	95.0	20	6	AX104844	Sequence
7	19	95.0	20	6	AX105120	Sequence
8	18.4	92.0	400	6	ARI64570	Sequence
9	18.4	92.0	697	9	S82168S1	IPF-1-insul
10	18.4	92.0	942	9	HSD35632	Human insul
11	18.4	92.0	992	9	HSD35632	Human insul
12	18.4	92.0	1428	9	HSD30329	Human insul
13	18.4	92.0	1525	9	HSGSFGENE	H.sapiens m
14	18.4	92.0	5658	6	ARI64589	Sequence
15	18.4	92.0	9117	6	AX345971	Sequence
16	18.4	92.0	32526	9	AL353195	Human DNA
17	18.4	92.0	208531	2	AC087560	Mus muscu
18	17.4	87.0	933	1	SGY08764	S.glaucosce
19	17.4	87.0	3595	8	ANI272133	Aspergill
20	17.4	87.0	25459	1	SGAJ6985	Streptomy
21	17.4	87.0	170696	2	AP004658	Oryza sat
22	17.4	87.0	179714	8	AP002743	Oryza sat
23	17	85.0	4668	2	AC014532	Drosophil
24	17	85.0	10369	1	AE005059	Halobacte
25	17	85.0	181438	3	AC008194	Drosophil
26	17	85.0	343807	3	AE003511	Drosophil
27	16.8	84.0	20	6	AX104882	Sequence
28	16.8	84.0	20	6	AX105261	Sequence
29	16.8	84.0	733	9	HSJ331096	Homo sapi
30	16.8	84.0	1481	10	MUSTNFR203	Mus muscu
31	16.8	84.0	1956	10	MUSTNFR2	Murine tumo
32	16.8	84.0	2048	10	MUSMTNFR1	Mouse tumor
33	16.8	84.0	2063	10	BMPS58	Murine mRN
34	16.8	84.0	2086	10	BC004599	Mus muscu
35	16.8	84.0	2154	10	MUSTNFX	Mus muscu
36	16.8	84.0	2179	10	MNTNFR5	Mouse mRN
37	16.8	84.0	2890	1	SGNATHRD	S.griseus g
38	16.8	84.0	35710	2	AC103128	Rattus no
39	16.8	84.0	39744	1	SC5H4	Streptomy
40	16.8	84.0	65672	2	AC100892	Mus muscu
41	16.8	84.0	73444	2	AC095922	Rattus no
42	16.8	84.0	89369	2	AC094673	Rattus no
43	16.8	84.0	90767	2	H0510A06	AL442104 Oryza sat
44	16.8	84.0	91160	2	AC105689	Rattus no
45	16.8	84.0	122064	2	OSJN00064	Oryza sat

ALIGNMENTS

RESULT	1	AX104880	Sequence	1072 from Patent WO0122972.	20 bp	DNA	linear	PAT 30-APR-2001
LOCUS	AX104880	Sequence	1072 from Patent WO0122972.					
DEFINITION	AX104880	Sequence	1072 from Patent WO0122972.					
ACCESSION	AX104880	Sequence	1072 from Patent WO0122972.					
VERSION	AX104880.1	GI:13921077						
KEYWORDS		synthetic construct.						
SOURCE		synthetic construct.						
ORGANISM		artificial sequence.						
REFERENCE		1 (bases 1 to 20)						
AUTHORS		Krieg, A.M., Schetter, C. and Vollmer, J.C.						
TITLE		Immunostimulatory nucleic acids						
JOURNAL		Patent: WO 0122972-A 1072 05-APR-2001;						
		UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical						
FEATURES		GmbH (DE)						
source		Location/Qualifiers						
		1..20						
		/organism="synthetic construct"						
		/db_xref="taxon:32630"						
BASE COUNT		2 a						
ORIGIN		13 g 2 t						

STREET: 30500 No. 6025183thwestern Highway, Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: U.S.
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/814,095
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 38,972
REFERENCE/DOCKET NUMBER: 2391.00066
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3096 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Alternatively spliced AchE
DESCRIPTION: comprising exons 2, 3, 4 and 5 as well as the translated portion
DESCRIPTION: of Intron 4 (readthrough)"
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 160..1959
US-08-814-095-3

Query Match 77.9%; Score 14.8; DB 3; Length 3096;
Best Local Similarity 88.9%; Pred. No. 2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtggg 18
||||||| |||||
Db 2973 GGGGACGTCGGGTGGG 2956

RESULT 15
US-08-976-255-2/c
Sequence 2, Application US/08976255
Patent No. 6136581
GENERAL INFORMATION:
APPLICANT: Jono, Keith E.
APPLICANT: Plowman, Gregory
TITLE OF INVENTION: KINASE GENES AND USES
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,255

FILING DATE: No. 6136581ember 21, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/031,675
FILING DATE: No. 6136581ember 22, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 229/182
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5267 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-976-255-2

Query Match 77.9%; Score 14.8; DB 3; Length 5267;
Best Local Similarity 88.9%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtggg 18
||||||| |||||
Db 1774 GGGGACGTCGACGGTGG 1757

Search completed: August 10, 2002, 03:06:30
Job time: 16056 sec

; ADDRESSEE: KOHN & ASSOC

;
 ; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM
 ;
 ; DATE OF INVENTION:
 ;
 ; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
 ;
 ; NUMBER OF SEQUENCES: 7
 ;
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: KOHN & ASSOCIATES
 ;
 ;

Gaps 0;

```
;
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,826A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 2391.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
; US-08-318-826A-7

Query Match 77.9%; Score 14.8; DB 2; Length 3016;
Best Local Similarity 88.9%; Pred. No. 2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

US-08-318-826A-7

Qy 1 ggggagctgcagctgggg 18
|||||
Db 2893 GGGGACGTCGGGTGGG 2876

RESULT 10
US-08-370-156-5/c
; Sequence 5, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
```

```
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
; US-08-370-156-5

Query Match 77.9%; Score 14.8; DB 2; Length 3016;
Best Local Similarity 88.9%; Pred. No. 2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ggggagctgcagctgggg 18
|||||
Db 2893 GGGGACGTCGGGTGGG 2876

RESULT 11
US-08-814-095-5/c
; Sequence 5, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Alternatively spliced ACHE
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
```

; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: /note= "splice variant: Exons 1, 2,
; OTHER INFORMATION: 3, 4 and 6"
US-08-318-826A-5

Query Match 77.9%; Score 14.8; DB 2; Length 2256;
Best Local Similarity 88.9%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 2

QY 1 ggggacgtcgactgggg 18
|||||
Db 2133 GGGACGTCGGGTGGG 2116

RESULT 7
US-08-370-156-1/c
; Sequence 1, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-370-156-1

Query Match 77.9%; Score 14.8; DB 2; Length 2256;
Best Local Similarity 88.9%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 2

QY 1 ggggacgtcgactgggg 18
|||||
Db 2133 GGGACGTCGGGTGGG 2116

RESULT 8
US-08-814-095-1/c
; Sequence 1, Application US/08814095

; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "ACHE gene comprising exons
; DESCRIPTION: 2, 3, 4 and 6"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-814-095-1

Query Match 77.9%; Score 14.8; DB 3; Length 2256;
Best Local Similarity 88.9%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 2

QY 1 ggggacgtcgactgggg 18
|||||
Db 2133 GGGACGTCGGGTGGG 2116

RESULT 9
US-08-318-826A-7/c
; Sequence 7, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan

;; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
;; NUMBER OF SEQUENCES: 27
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Reising, Ethington, Barnard & Perry
;; STREET: P.O. Box 4390
;; CITY: Troy
;; STATE: Michigan
;; COUNTRY: US
;; ZIP: 48099
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/370,156
;; FILING DATE:
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kohn, Kenneth I.
;; REGISTRATION NUMBER: 30,955
;; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (810) 689-3500
;; TELEFAX: (810) 689-4071
;; INFORMATION FOR SEQ ID NO: 26:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1215 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..78
US-08-370-156-26

Query Match 77.9%; Score 14.8; DB 2; Length 1215;
Best Local Similarity 88.9%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 18
||||||| |||||
Db 1092 GGGGACGTCGGGTGGG 1075

RESULT 5
US-08-964-652-1
;; Sequence 1, Application US/08964652
;; Patent No. 6180387
;; GENERAL INFORMATION:
;; APPLICANT: Burnham, Martin K.R.
;; APPLICANT: Lonetto, Michael A.
;; APPLICANT: Warren, Patrick V.
;; APPLICANT: Biswas, Sanjoy
;; APPLICANT: Warren, Richard L.
;; TITLE OF INVENTION: NOVEL ARGININE DEIMINASE
;; NUMBER OF SEQUENCES: 7
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dechert Price & Rhoads
;; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
;; CITY: Philadelphia
;; STATE: PA
;; COUNTRY: US
;; ZIP: 19103
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/964,652
;; FILING DATE:

;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Dickinson, Todd Q
;; REGISTRATION NUMBER: 28,354
;; REFERENCE/DOCKET NUMBER: GM10056
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 215-994-2252
;; TELEFAX: 215-994-2222
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1236 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
US-08-964-652-1

Query Match 77.9%; Score 14.8; DB 4; Length 1236;
Best Local Similarity 88.9%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gggacgtcgacgtgggg 19
||||| |||||
Db 1176 GGTACTGGACGTGGGG 1193

RESULT 6
US-08-318-826A-5/C
;; Sequence 5, Application US/08318826A
;; Patent No. 5891725
;; GENERAL INFORMATION:
;; APPLICANT: Soreq, Hermona
;; APPLICANT: Zakut, Haim
;; APPLICANT: Eckstein, Fritz
;; TITLE OF INVENTION: Synthetic Antisense
;; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
;; TITLE OF INVENTION: Containing Them
;; NUMBER OF SEQUENCES: 9
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Kohn & Associates
;; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
;; CITY: Farmington Hills
;; STATE: Michigan
;; COUNTRY: US
;; ZIP: 48334
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/318,826A
;; FILING DATE:
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kohn, Kenneth I.
;; REGISTRATION NUMBER: 30,955
;; REFERENCE/DOCKET NUMBER: 2391.00001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (248) 539-5050
;; TELEFAX: (248) 539-5055
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2256 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA to mRNA
;; HYPOTHETICAL: NO

NAME/KEY: CDS
LOCATION: 31329..36071
FEATURE:
NAME/KEY: CDS
LOCATION: 36155..41830
US-08-804-227C-7

Query Match 81.1%; Score 15.4; DB 2; Length 44377;
Best Local Similarity 94.1%; Pred. No. 85;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ggacgtcgcacgtg9ggg 19
|||||
Db 18213 GGAGTCGACGTCGGG 18197

RESULT 2
US-08-804-198-1/c
; Sequence 1, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rostock, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830

US-08-804-198-1

Query Match 81.1%; Score 15.4; DB 2; Length 44377;
Best Local Similarity 94.1%; Pred. No. 85;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ggacgtcgcacgtg9ggg 19
|||||
Db 18213 GGAGTCGACGTCGGG 18197

RESULT 3
US-08-370-156-24/c
; Sequence 24, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-370-156-24

Query Match 77.9%; Score 14.8; DB 2; Length 374;
Best Local Similarity 88.9%; Pred. No. 2.5e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggacgtcgcacgtg9ggg 18
|||||
Db 251 GGGGACGTCGGGTGGG 234

RESULT 4
US-08-370-156-26/c
; Sequence 26, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:06:24 ; Search time 277.54 Seconds
(without alignments)
16.816 Million cell updates/sec

Title: US-09-672-126-30
Perfect score: 19
Sequence: 1 ggggacgtgcacgtgggg 19

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	15.4	81.1	44377	2	US-08-804-227C-7
C 2	15.4	81.1	44377	2	US-08-804-198-1
C 3	14.8	77.9	374	2	US-08-370-156-24
C 4	14.8	77.9	1215	2	US-08-370-156-26
C 5	14.8	77.9	1236	4	US-08-964-652-1
C 6	14.8	77.9	2256	2	US-08-318-826A-5
C 7	14.8	77.9	2256	2	US-08-370-156-1
C 8	14.8	77.9	2256	3	US-08-814-095-1
C 9	14.8	77.9	3016	2	US-08-318-826A-7
C 10	14.8	77.9	3016	2	US-08-370-156-5
C 11	14.8	77.9	3016	3	US-08-814-095-5
C 12	14.8	77.9	3096	2	US-08-318-826A-6
C 13	14.8	77.9	3096	2	US-08-370-156-3
C 14	14.8	77.9	3096	3	US-08-814-095-3
C 15	14.8	77.9	5267	3	US-08-976-255-2
C 16	14.8	77.9	12394	4	US-09-488-856A-10
C 17	14.8	77.9	35060	3	US-08-814-095-7
C 18	14.4	75.8	10763	1	US-08-761-258-1
C 19	14.4	75.8	10763	2	US-08-977-306-1
C 20	14.4	75.8	13987	2	US-08-804-227C-13
C 21	14.4	75.8	35060	3	US-08-814-095-7
C 22	14.4	75.8	44377	2	US-08-804-227C-7
C 23	14.4	75.8	44377	2	US-08-804-198-1
C 24	14.4	75.8	4403765	4	US-09-103-840A-2
C 25	14.4	75.8	4411529	4	US-09-103-840A-1
C 26	14.2	74.7	87	4	US-08-936-477-7
C 27	14.2	74.7	143	4	US-09-025-769B-263

C 28	14.2	74.7	400	4	US-08-881-450A-1
C 29	14.2	74.7	420	1	US-08-470-179-108
C 30	14.2	74.7	756	4	US-08-413-974-3
C 31	14.2	74.7	756	4	US-08-434-418-3
C 32	14.2	74.7	756	4	US-08-433-288-3
C 33	14.2	74.7	756	4	US-08-174-739A-3
C 34	14.2	74.7	877	3	US-09-165-240-3
C 35	14.2	74.7	877	4	US-09-568-059-3
C 36	14.2	74.7	950	4	US-09-230-421-1
C 37	14.2	74.7	1107	2	US-08-933-750C-77
C 38	14.2	74.7	1107	3	US-09-234-613-77
C 39	14.2	74.7	1272	2	US-09-068-109-1
C 40	14.2	74.7	1518	1	US-08-660-765A-1
C 41	14.2	74.7	1593	4	US-08-793-044-2
C 42	14.2	74.7	1947	4	US-09-025-769B-264
C 43	14.2	74.7	2051	1	US-08-343-785-7
C 44	14.2	74.7	2051	2	US-08-462-221-7
C 45	14.2	74.7	2051	3	US-08-946-458-7

ALIGNMENTS

RESULT 1
US-08-804-227C-7/c
; Sequence 7, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:

Query Match 83.2%; Score 15.8; DB 9; Length 212;
Best Local Similarity 89.5%; Pred. No. 1.1e+04;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
|||||
Db 63 GGGGATGTGGACGTGGGG 81

Search completed: August 10, 2002, 02:11:23
Job time: 13144 sec

Query Match 83.2%; Score 15.8; DB 9; Length 190;
 Best Local Similarity 89.5%; Pred. No. 1.1e+04;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
 ||||| ||||| ||||| |||||
 Db 152 GGGGAGCTCCCGTGGGG 170

RESULT 14
 BH407485
 LOCUS 1007005D04.x1 1007 - RescueMu Grid H Zea mays genomic, DNA
 DEFINITION
 ACCESSION BH407485 201 bp DNA linear GSS 12-DEC-2001
 VERSION BH407485
 KEYWORDS BH407485.1 GI:17572454
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC.
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
 AUTHORS Walbot,V.
 TITLE Maize genomic sequences found using engineered RescueMu transposon
 JOURNAL Unpublished (2001)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu

Possible ligation site so sequence was trimmed. Post-ligation
 sequence submitted separately.
 Plate: 1007005 column: 2
 Class: transposon-tagged.

FEATURES
 source
 1..201
 /organism="Zea mays"
 /cultivar="mixed background W23/A188/B73"
 /db_xref="taxon:4577"
 /clone_lib="1007 - RescueMu Grid H"
 /tissue_type="leaf"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site_1: BamHI; Site_2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid H was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 36 a 46 c 88 g 31 t

Query Match 83.2%; Score 15.8; DB 12; Length 201;
 Best Local Similarity 89.5%; Pred. No. 1.1e+04;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
 ||||| ||||| ||||| |||||
 Db 77 GGGGAGCTCCCGTGGGAG 95

RESULT 15
 BB601910

LOCUS BB601910 212 bp mRNA linear EST 05-DEC-2000
 DEFINITION BB601910 RIKEN full-length enriched, 13 days embryo lung Mus musculus cDNA clone D430015N04 5', mRNA sequence.
 ACCESSION BB601910
 VERSION BB601910.1 GI:11553312
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 212)
 Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P., Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T., Hodojima,Y., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J., Kojima,Y., Konno,H., Kusakabe,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Okazaki,Y., Okido,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Sato,K., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka,T., Toya,T., Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshiki,A., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Aizawa, K. et al. 2000)
 Unpublished (2000)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-5222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES
 source

Location/Qualifiers
 1..212
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="D430015N04"
 /clone_lib="RIKEN full-length enriched, 13 days embryo lung"
 /tissue_type="lung"
 /dev_stage="13 days embryo"
 /lab_host="DH10B"
 /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
 GAGAGAGAGCGCCGACCTCGAGTTTTTTTTTTTTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'
 GAGAGAGAGATTCTCGAGTTAATAATATATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
 40 a 42 c 91 g 39 t

BASE COUNT
 ORIGIN

AUTHORS

Cordonnier-Pratt, M.-M., Gingle, A., Dean, R., Sudman, M. and Pratt, L.H.
 An EST database from Sorghum: pathogen-induced plants
 Unpublished (2000)
 Contact: Cordonnier-Pratt MM
 Department of Botany
 The University of Georgia
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 542 1805
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude polyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
 Seq primer: JEN REV
 High quality sequence stop: 564
 POLYA=No.

FEATURES

source

1. .570
 Location/Qualifiers
 /organism="Sorghum bicolor"
 /db_xref="taxon:4538"
 /clone_lib="Pathogen induced 1 (PI1)"
 /note="Organ: Anthracnose-infected leaves from two-week-old sorghum plants 48 hr after inoculation; Vector: pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2: EcoRI; Two-week-old sorghum plants (BX 623 cultivar) were infected with pathogen (isolate FRM421 of Colletotrichum graminicola, which is a sorghum isolate). RNA was prepared from infected leaves harvested from 45 seedlings 48 hours after inoculation. Note: young seedlings (2 weeks old) exhibit juvenile resistant reaction, which is an incompatible interaction. As they grow older (4 weeks or older), plants resume susceptibility to anthracnose disease. The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."

BASE COUNT
 ORIGIN

137 a 138 c 165 g 130 t

Query Match

Best Local Similarity 84.2%; Score 16; DB 10; Length 570;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgg 16

|||||

Db 145 GGGGACGTCGACGTGG 130

RESULT 12
 BF569128/c

LOCUS 602184525T1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4300327 3',
 DEFINITION mRNA sequence. EST 12-DEC-2000

ACCESSION BF569128

VERSION BF569128.1 GI:11642508

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1224)

NH-MGC http://mhc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-re@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LUCM159 row: b column: 08

High quality sequence start: 23

High quality sequence stop: 711.

FEATURES

source

1. .1224
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4300327"
 /clone_lib="NIH_MGC_42"
 /tissue_type="epithelioid carcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. |"
 BASE COUNT 366 a 296 c 364 g 197 t
 ORIGIN

Query Match 84.2%; Score 16; DB 10; Length 1224;

Best Local Similarity 100.0%; Pred. No. 1.le-04;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgg 16

|||||

Db 1079 GGGGACGTCGACGTGG 1064

RESULT 13

AJ399447

LOCUS

DEFINITION AJ399447 dkfz426 Gallus gallus cDNA clone 9p12r1, mRNA sequence. EST 25-JAN-2001

ACCESSION AJ399447

VERSION AJ399447.1 GI:7134431

KEYWORDS EST.

SOURCE chicken.

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 190)

AUTHORS Abdrakhmanov, I., Lodygin, D., Geroth, P., Arakawa, H., Law, A., Plachy, J., Korn, B. and Buerstedde, J.M.

TITLE A large database of chicken bursal ESTs as a resource for the

analysis of vertebrate gene function

JOURNAL Genome Res. 10 (12), 2062-2069 (2000)

MEDLINE 20568495

COMMENT Contact: Buerstedde JM

Cellular Immunology

Heinrich-Pette-Institute

Martinistr. 52, 20251 Hamburg, Germany

Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

Location/Qualifiers

source

1. .190

/organism="Gallus gallus"

/strain="CB"

/db_xref="taxon:9031"

/clone="9p12r1"

/clone_lib="dkfz426"

/tissue_type="Bursa of Fabricius"

BASE COUNT 23 a 72 c 70 g 25 t

ORIGIN

```

/db_xref="taxon:99883"
/clone="020124"
/clone_lib="G"
/note="Genoscope sequence ID : C0B020BE12L1P1-end : T7"
BASE COUNT 186 a 267 c 303 g 173 t 8 others
ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 937;
Best Local Similarity 94.4%; Pred. No. 7.9e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ggggacgtcgactggg 18
|||||
Db 864 GGGAGCTCGCGTGGG 881

RESULT 9
AQ448518/c
LOCUS
DEFINITION mgxb0020L10f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
clone mgxb0020L10f, DNA sequence.
ACCESSION AQ448518
VERSION AQ448518.1 GI:4577655
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
1 (bases 1 to 464)
Yu.Y., Zhu,H., Boyd,C.A., Gaudette,B., Gayle,A., Kingsbury,R.,
Phillips,K., Sasinowski,M, Wing,R.A. and Dean,R.A.
A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence stop: 383.
FEATURES
Location/Qualifiers
1..464
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0020L10f"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 143 a 107 c 111 g 103 t
ORIGIN

Query Match 84.2%; Score 16; DB 12; Length 464;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggggacgtcgactggg 16
|||||
Db 151 GGGAGCTCGACGTGG 136

RESULT 11
BE594603/c
LOCUS
DEFINITION P11_35_C12.bl_A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,
mRNA sequence.
ACCESSION BE594603
VERSION BE594603.1 GI:9849676
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 570)

```

```

Qy 2 ggggacgtcgactggg 17
|||||
Db 289 GGGAGCTCGACGTGG 274

RESULT 10
AW680752/c
LOCUS
DEFINITION WSL_7_A06.bl_A002 Water-stressed 1 (WS1) Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION AW680752
VERSION AW680752.1 GI:7554553
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 470)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
,L.H.
An EST database from Sorghum: water-stressed plants
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 438
POLYA-No.
Location/Qualifiers
1..470
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Water-stressed 1 (WS1)"
/note="Organ: Mix of 5-week old plants on days 7 & 8 after
water was withheld; Vector: Lambda zap; Site_1: XhoI;
Site_2: EcoRI; The library was made from poly-A RNA in the
cloning vector lambda zap II. Clones to be sequenced were
prepared by mass excision."
BASE COUNT 111 a 115 c 146 g 98 t
ORIGIN

Query Match 84.2%; Score 16; DB 9; Length 470;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggggacgtcgactggg 16
|||||
Db 151 GGGAGCTCGACGTGG 136

RESULT 11
BE594603/c
LOCUS
DEFINITION P11_35_C12.bl_A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,
mRNA sequence.
ACCESSION BE594603
VERSION BE594603.1 GI:9849676
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 570)

```

```

/clone="UUGC1M0437N12"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydronamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gii4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      212 a 168 c 127 g 204 t
ORIGIN

Query Match      86.3%; Score 16.4; DB 12; Length 711;
Best Local Similarity 94.4%; Pred. No. 7.7e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gggagctgcagctgggg 19
||||| |||||||
Db 377 GGGACGTTGACGTGGGG 360

RESULT 7
LOCUS      AQ288578/c
DEFINITION nbxb0033118f CUGI Rice BAC Library Oryza sativa genomic clone
VERSION     AQ288578
KEYWORDS    AQ288578.1 GI:3950192
SOURCE      GSS.
ORGANISM    Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Lillipsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE   1 (bases 1 to 754)
            Wing,R.A. and Dean,R.A.
            A BAC End Sequencing Framework to Sequence the Rice Genome
            Unpublished (1998)
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TATACGACTACTATAGG
            Class: BAC ends
            High quality sequence start: 3
            High quality sequence stop: 58.
            Location/Qualifiers
                1..754
                /organism="Oryza sativa"
                /strain="Japonica"
                /cultivar="Nipponbare"
                /db_xref="taxon:4530"
                /clone_lib="nbxb0033118f"
                /clone_lib="CUGI Rice BAC Library"
                /tissue_type="Leaf"

```

```

/lab_host="E. coli DH10B"
/notes="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."
BASE COUNT      71 a 347 c 138 g 194 t
ORIGIN

Query Match      86.3%; Score 16.4; DB 12; Length 754;
Best Local Similarity 94.4%; Pred. No. 7.7e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggagctgcagctgggg 18
||||| |||||||
Db 263 GGGCCGCTGACGTGGGG 246

RESULT 8
LOCUS      CNS03E1W
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
VERSION     AL240449.1 GI:7961218
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
ORGANISM    Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1 (bases 1 to 937)
            Roest-Crolius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
            Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
            Unpublished
            2 (bases 1 to 937)
            Roest-Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
            Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
            Unpublished
            REFERENCE 3 (bases 1 to 937)
            Genoscope.
            Direct Submission
            Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
            This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
            Location/Qualifiers
                1..937
                /organism="Tetraodon nigroviridis"

```


Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gggacgtcgacgtggggg 19
 DB 235 GGGACGTGACGTGGGG 218

RESULT 2

AI548322 510 bp mRNA linear EST 22-MAR-1999
 LOCUS UI-R-C3-tg-f-08-0-UI.s1 UI-R-C3 Rattus norvegicus cDNA clone
 DEFINITION UI-R-C3-tg-f-08-0-UI 3', mRNA sequence.

ACCESSION AI548322
 VERSION AI548322.1 GI:4465810
 KEYWORDS EST.
 SOURCE Norway rat.

ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 510)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT

Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized brain library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 466-500, >AT-rich/Low_complexity

Seq primer: M3 Forward.

FEATURES

Location/Qualifiers

1..510
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C3-tg-f-08-0-UI"
 /dev_stage="adult"
 /lab_host="UI-R-C3"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site1: Not I; Site2: Eco RI; The UI-R-C3 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day embryos, after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1, UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C3) was constructed as follows: PCR amplified cDNA inserts from UI-R-C2p clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C2p library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and

electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C3 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)"
 BASE COUNT 153 a 118 c 93 g 146 t
 ORIGIN

Query Match 86.3%; Score 16.4; DB 9; Length 510;
 Best Local Similarity 94.4%; Pred. No. 7.4e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gggacgtcgacgtggggg 19
 DB 210 GGGACGTGACGTGGGG 193

RESULT 3

CNS04KVB 568 bp DNA linear GSS 21-MAY-2000
 LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
 DEFINITION 117C12 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION AL295328
 VERSION AL295328.1 GI:8033908
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 568)

AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach,J.

TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 568)

AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.

TITLE Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 568)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases

COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.

FEATURES

Location/Qualifiers

1..568
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="117C12"
 /clone_lib="G"

BASE COUNT 128 a 154 c 186 g 94 t 6 others
 ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 568;
 Best Local Similarity 94.4%; Pred. No. 7.5e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggacgtcgacgtgggg 18
 DB 232 GGGACGTGCGCGTGGGG 215

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:11:20 ; Search time 9068.22 seconds
(without alignments)
28.279 Million cell updates/sec

Title: US-09-672-126-30

Perfect score: 19

Sequence: 1 ggggacgtgcacgtggggg 19

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estt:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.4	86.3	473	9	AI229090
2	16.4	86.3	510	9	AI548322
3	16.4	86.3	568	12	CNS04KVB
4	16.4	86.3	611	9	AI622583
5	16.4	86.3	662	12	AG040749
6	16.4	86.3	711	12	AZ611390
7	16.4	86.3	754	12	AQ288578
8	16.4	86.3	937	12	CNS03E1W
9	16.4	84.2	464	12	AQ448518
10	16.4	84.2	470	9	AM680752
11	16.4	84.2	570	10	BE594603
12	16.4	84.2	1224	10	BF569128
13	15.8	83.2	190	9	AJ399447
14	15.8	83.2	201	12	BH407485
15	15.8	83.2	212	9	BB601910
16	15.8	83.2	244	9	AV055131
17	15.8	83.2	258	9	BB216437

c 18	15.8	83.2	264	10	BG463735
19	15.8	83.2	278	10	BG609203
20	15.8	83.2	302	10	BE598895
21	15.8	83.2	317	10	BG241925
22	15.8	83.2	324	9	AI113135
c 23	15.8	83.2	330	10	BI795387
24	15.8	83.2	331	10	BG644265
25	15.8	83.2	367	10	BG159065
c 26	15.8	83.2	377	10	BF358668
27	15.8	83.2	397	10	BF201606
c 28	15.8	83.2	398	10	BM158468
29	15.8	83.2	417	10	BI338710
30	15.8	83.2	418	10	BI338705
c 31	15.8	83.2	433	10	BF517606
32	15.8	83.2	440	9	AA533812
33	15.8	83.2	454	9	BE051161
34	15.8	83.2	460	12	AQ912924
35	15.8	83.2	464	9	AU070161
36	15.8	83.2	472	10	BE500778
c 37	15.8	83.2	475	10	BE358201
c 38	15.8	83.2	483	10	BF516406
39	15.8	83.2	491	9	AW918937
c 40	15.8	83.2	492	12	FR0040043
41	15.8	83.2	494	10	BM427459
c 42	15.8	83.2	518	10	BF625803
43	15.8	83.2	525	10	BJ201959
c 44	15.8	83.2	529	10	BE586834
45	15.8	83.2	536	10	BI339722

ALIGNMENTS

RESULT 1
AI229090/c
LOCUS 473 bp mRNA linear EST 30-OCT-1998
DEFINITION EST225785 Normalized rat brain, Bento Soares Rattus sp. cDNA clone
RBRDD85 3' end, mRNA sequence.
ACCESSION AI229090.1 GI:3812977
VERSION AI229090.1
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 473)
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.
TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.
FEATURES
Location/Qualifiers
source 1. 473
/organism="Rattus sp."
/db_xref="taxon:10118"
/clone="RBRDD85"
/clone_lib="Normalized rat brain, Bento Soares"
/note="Organ: brain; Vector: pT73pac; Site_1: EcoRI; Site_2: NotI"
BASE COUNT 145 a 122 c 92 g 114 t
ORIGIN
Query Match 86.3%; Score 16.4; DB 9; Length 473;
Best Local Similarity 94.4%; Pred. NO. 7.4e+03;

BG463735 EML_51_D0
BG609203 322764 MA
BE598895 P11_83_B0
BG241925 RH122_17_
AI113135 UI-R-C2p-
BI795387 HO19E06 E
BG644265 MO1m39 M0
BG159065 RH122_17_
BF358668 OVI-ET000
BF201606 WHEI772_E
BM158468 NXLV_034_
BI338710 362574 MA
BI338705 362568 MA
BF517606 NXSL_026_
AA533812 nj94a08.S
BE051161 za72H04.b
AQ912924 nbeb0039C
AU070161 AU070161
BE500778 WHE0991-0
BE358201 DGL_26_H0
BF516406 UI-H-BW1-
AW918937 EST350241
AI127537 Fugu rubr
BM427459 pgt2n.pk0
BF625803 HVSMEa001
BJ201959 BJ201959
BE586834 WHE0508_F
BI339722 364785 MA

THIS PAGE BLANK (USPTO)

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 228 BP; 62 A; 64 C; 72 G; 30 T; 0 other;

Query Match 83.2%; Score 15.8; DB 23; Length 228;
Best Local Similarity 89.5%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
||||| ||||| ||||| |||||
Db 98 ggggaggtcgtcgtgggg 116

RESULT 15

ABA09482
ID ABA09482 standard; cDNA; 1974 BP.

XX ABA09482;

XX
DT 11-JAN-2002 (first entry)

DE Human secreted protein homologue-encoding cDNA, SEQ ID NO:1258.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative-retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteoparathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnary; antiulcer; ss.

OS Homo sapiens.

XX WO200157188-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US03800.

XX 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457740/49.

DR P-PSDB; ABB12238.

XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX

PS Claim 1; Page 962-963; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence

CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.

CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.

XX SQ Sequence 1974 BP; 305 A; 657 C; 640 G; 372 T; 0 other;

Query Match 83.2%; Score 15.8; DB 22; Length 1974;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
||||| ||||| ||||| |||||

Db 1313 ggggcccgtcgtcgtgggg 1331

Search completed: August 10, 2002, 03:21:54

Job time: 13685 sec

PR 27-SEP-1999; 99US-0156147.
 XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX Hartmann G, Bratzler RL, Krieg A;
 XX WPI; 2001-290487/30.
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 XX Claim 201; Page 103; 168pp; English.
 XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX Sequence 21 BP; 2 A; 3 C; 14 G; 2 T; 0 other;
 SQ

Query Match 83.2%; Score 15.8; DB 22; Length 21;
 Best Local Similarity 89.5%; Pred. No. 6.3e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ggggacgtcgactgtgggg 19
 ||||| |||||
 Db 1 ggggacgtcgactgtgggg 19

RESULT 13
 ID AAF99873
 XX AAF99873 standard; DNA; 21 BP.
 XX AAF99873;
 XX 12-JUN-2001 (first entry)
 DT Immunostimulatory nucleic acid #989.
 DE Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 XX immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 OS Synthetic.
 XX WO200122972-A2.
 PN 05-APR-2001.
 PD 25-SEP-2000; 2000WO-US26383.
 PF 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 XX Krieg AM, Schetter C, Vollmer J;
 XX WPI; 2001-273485/28.
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 XX

PS Claim 101; Page 59; 338pp; English.
 XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 XX Sequence 21 BP; 2 A; 3 C; 14 G; 2 T; 0 other;
 SQ

Query Match 83.2%; Score 15.8; DB 22; Length 21;
 Best Local Similarity 89.5%; Pred. No. 6.3e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ggggacgtcgactgtgggg 19
 ||||| |||||
 Db 1 ggggacgtcgactgtgggg 19

RESULT 14
 ID ABL25165
 XX ABL25165 standard; DNA; 228 BP.
 XX ABL25165;
 XX 26-MAR-2002 (first entry)
 DT Drosophila melanogaster genomic polynucleotide SEQ ID NO 26968.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 KW Drosophila melanogaster.
 OS WO200171042-A2.
 PN 27-SEP-2001.
 PD 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PR (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Claim 1; SEQ ID NO 26968; 21pp + Sequence Listing; English.
 PS The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

```

PN WO200122972-A2.
XX
PD 05-APR-2001.
XX
PF 25-SEP-2000; 2000WO-US26383.
XX
PR 25-SEP-1999; 99US-0156113.
XX
PR 27-SEP-1999; 99US-0156135.
XX
PR 23-AUG-2000; 2000US-0227436.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
XX
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX
XX
DR WPI; 2001-273485/28.
XX
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 57; 338pp; English.
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SQ Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 83.2%; Score 15.8; DB 22; Length 20;
Best Local Similarity 89.5%; Pred. No. 6.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtggggg 19
Db 1 ggggtcgtcgacgaggggg 19

RESULT 11
AAF99830
ID AAF99830 standard; DNA; 20 BP.
XX
AC AAF99830;
XX
DT 12-JUN-2001 (first entry)
XX
DE Immunostimulatory nucleic acid #946.
XX
KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
OS Synthetic.
XX
PN WO200122972-A2.
XX
PD 05-APR-2001.
XX
PF 25-SEP-2000; 2000WO-US26383.
XX
PR 25-SEP-1999; 99US-0156113.
XX
PR 27-SEP-1999; 99US-0156135.
XX
PR 23-AUG-2000; 2000US-0227436.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
XX
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX
XX
DR WPI; 2001-273485/28.
XX
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 57; 338pp; English.
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SQ Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

```

```

PR 27-SEP-1999; 99US-0156135.
PR 23-AUG-2000; 2000US-0227436.
XX
XX
PA (IOWA ) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX
XX
DR WPI; 2001-273485/28.
XX
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
XX
PS Claim 101; Page 58; 338pp; English.
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SQ Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 83.2%; Score 15.8; DB 22; Length 20;
Best Local Similarity 89.5%; Pred. No. 6.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtggggg 19
Db 1 ggggtcgtcgacgaggggg 19

RESULT 12
AAF98767
ID AAF98767 standard; DNA; 21 BP.
XX
XX
AC AAF98767;
XX
DT 11-JUN-2001 (first entry)
XX
DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 37.
XX
KW Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..2
FT /*tag= a
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
FT modified_base 16..20
FT /*tag= b
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
XX
PN WO200122990-A2.
XX
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26527.
XX
XX

```

KW metabolic pathway engineering; catabolic pathway engineering; ss.
 XX Aspergillus oryzae.
 OS
 PN WO200056762-A2.
 XX
 PD 28-SEP-2000.
 XX
 XX 22-MAR-2000; 2000WO-US07781.
 XX
 XX 22-MAR-1999; 99US-0273623.
 XX
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 XX
 XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
 PI WPI; 2000-594572/56.
 DR
 XX Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags
 XX
 PS Claim 88; Page 2020; 3161pp; English.
 XX
 CC The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring
 CC the global expression of genes from FF cells allows the production
 CC potential of the microorganisms to be improved. New genes may be
 CC discovered, possible functions of unknown open reading frames can be
 CC identified and gene copy number variation and stability can be
 CC monitored. The expression of genes can be used to study how FF cells
 CC adapt to changes in culture conditions, environmental stress, spore
 CC morphogenesis, recombination, metabolic or catabolic pathway
 CC engineering. Using ESTs provides several advantages over genomic or
 CC random cDNA clones including elimination of redundancy as one spot on an
 CC array equals one gene or open reading frame, and organisation of the
 CC microarrays based on function of the gene products to facilitate
 CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
 CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
 CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
 CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
 CC all specifically claimed in the present invention.
 XX
 SQ Sequence 701 BP; 154 A; 215 C; 192 G; 140 T; 0 other;

Query Match 84.2%; Score 16; DB 21; Length 701;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gggagcgtcagctgg 16
 |||||
 DB 400 GGGAGCGTCGCGTGG 385

RESULT 9
 AAF98748
 ID AAF98748 standard; DNA; 20 BP.
 XX
 XX AAF98748;
 AC
 XX
 XX 11-JUN-2001 (first entry)
 DT
 XX Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 18.
 DE
 XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
 KW

KW viral infection; phosphorothioate backbone; palindromic; cancer; ds.
 XX Synthetic.
 OS
 PH Location/Qualifiers
 FT modified_base 1..2
 FT /*tag= a
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 FT 15..19
 FT /*tag= b
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 XX
 XX WO200122990-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 27-SEP-2000; 2000WO-US26527.
 XX
 XX 27-SEP-1999; 99US-0156147.
 XX
 XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX
 XX Hartmann G, Bratzler RL, Krieg A;
 PI WPI; 2001-290487/30.
 DR
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid
 XX
 PS Claim 201; Page 103; 168pp; English.
 XX
 CC The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX
 SQ Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 83.2%; Score 15.8; DB 22; Length 20;
 Best Local Similarity 89.5%; Pred. No. 6.3e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gggagcgtcagctggggg 19
 |||||
 DB 1 gggggtcgtcagcagggggg 19

RESULT 10
 AAF99768
 ID AAF99768 standard; DNA; 20 BP.
 XX
 XX AAF99768;
 AC
 XX
 XX 12-JUN-2001 (first entry)
 DT
 XX Immunostimulatory nucleic acid #884.
 DE
 XX
 XX Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX
 OS Synthetic.
 XX

QY 1 ggggacgtcgcacgtgggg 19
 Db 1 ggggacgtcgcgtgggg 19

RESULT 6

AAF99870
 ID AAF99870 standard; DNA; 20 BP.

AC AAF99870;
 XX
 DT 12-JUN-2001 (first entry)
 XX
 DE Immunostimulatory nucleic acid #986.
 XX
 KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.

OS Synthetic.

PN WO200122972-A2.

PD 05-APR-2001.

PF 25-SEP-2000; 2000WO-US26383.

PR 25-SEP-1999; 99US-0156113.

PR 27-SEP-1999; 99US-0156135.

PR 23-AUG-2000; 2000US-0227436.

XX (IOWA) UNIV IOWA RES FOUND.

PA (COLE-) COLEY PHARM GMBH.

XX Krieg AM, Schetter C, Vollmer J;

DR WPI; 2001-273485/28.

XX Vaccinating against tumors, infectious diseases, allergies and asthma,
 PT using immunostimulatory Py-rich and TG nucleic acids -
 XX
 PS Claim 101; Page 59; 338pp; English.

XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.

XX Sequence 20 BP; 1 A; 3 C; 13 G; 3 T; 0 other;

Query Match 91.6%; Score 17.4; DB 22; Length 20;

Best Local Similarity 94.7%; Pred. No. 1.3e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgtcgcacgtgggg 19

Db 1 ggggacgtcgcgtgggg 19

RESULT 7

AAQ20217

ID AAQ20217 standard; DNA; 1170 BP.

XX AAQ20217;

XX 15-APR-1992 (first entry)

DE Sequence of tuf3 gene encoding translation elongation factor Tu3.

XX Elfamycin resistant actinomycetes; antibiotic resistant;

KW elongation factor; ss.

XX Streptomyces ramocissimus.

XX Key Location/Qualifiers

FT CDS 4..1170

XX /*tag= a

PN EP466251-A.

XX 15-JAN-1992.

XX 02-JUL-1991; 91EP-0201702.

XX 02-JUL-1991; 91EP-0201702.

PR 10-JUL-1990; 90EP-0201851.

XX (KONN) GIST-BROCADES NV.

XX Luiiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;

PI Woudt LP;

XX WPI; 1992-017874/03.

DR P-PSDB; AAR20244.

XX New protein conferring resistance to elfamycin - used to

PT transform streptomycetes to resistant pheno-type

XX Example; Pages 19-21; 35pp; English.

XX Substitution of residue 378 of the elongation factor (EF-Tu) with a
 CC valine, threonine, proline or phenylalanine results in an elfamycin
 CC resistant protein (EP-TuR). The advantage of this change is that
 CC the limiting factor for the prodn. of elfamycin by actinomycetes is
 CC removed by mutating the gene tuf into tufR encoding a protein
 CC resistant to an elfamycin, pref. mocimycin (Kirmomycin). The
 CC inventors claim EF-TuR and the genes (tufR) encoding it.

XX Sequence 1170 BP; 177 A; 378 C; 437 G; 178 T; 0 other;

Query Match 91.6%; Score 17.4; DB 13; Length 1170;

Best Local Similarity 94.7%; Pred. No. 90;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgtcgcacgtgggg 19

Db 1009 ggggacgtcgcacgtgggg 1027

RESULT 8

AAF12250/C

ID AAF12250 standard; cDNA; 701 BP.

XX AAF12250;

XX 13-MAR-2001 (first entry)

DE Aspergillus oryzae EST SEQ ID NO:4773.

XX Multiple gene expression; filamentous fungal cell; EST;

KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;

KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;

KW culture condition; environmental stress; spore morphogenesis;

CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX
SQ Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19

Db 1 ggggacgtcgacgtgggg 19

RESULT 4

AAF99767
ID AAF99767 standard; DNA; 20 BP.

XX AAF99767;

DT 12-JUN-2001 (first entry)

DE Immunostimulatory nucleic acid #883.

XX Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS Synthetic.

XX WO200122972-A2.

PN 05-APR-2001.

XX 25-SEP-2000; 2000WO-US26383.

XX 25-SEP-1999; 99US-0156113.

PR 27-SEP-1999; 99US-0156135.

PR 23-AUG-2000; 2000US-0227436.

XX (IOWA) UNIV IOWA RES FOUND.

PA (COLE-) COLEY PHARM GMBH.

XX Krieg AM, Schetter C, Vollmer J;

XX WPI; 2001-273485/28.

XX Vaccinating against tumors, infectious diseases, allergies and asthma

PT using immunostimulatory Py-rich and TG nucleic acids -

XX Claim 101; Page 57; 338pp; English.

XX The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.

XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19

Db 1 ggggacgtcgacgtgggg 19

RESULT 5

AAF98880
ID AAF98880 standard; DNA; 20 BP.

XX AAF98880;

DT 11-JUN-2001 (first entry)

DE Immunostimulatory nucleic acid assay control oligo SEQ ID NO: 161.

XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindromic; cancer; ds.
XX
OS Synthetic.

XX Key Location/Qualifiers

FT modified_base 1..2

FT FT /*tag= a

FT FT /mod_base= "OTHER"

FT FT /note= "phosphorothioate linkage"

FT modified_base 15..19

FT FT /*tag= b

FT FT /mod_base= "OTHER"

FT FT /note= "phosphorothioate linkage"

XX WO200122990-A2.

XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26527.

XX 27-SEP-1999; 99US-0156147.

XX (COLE-) COLEY PHARM GROUP INC.

PA (IOWA) UNIV IOWA RES FOUND.

XX Hartmann G, Bratzler RL, Krieg A;

XX WPI; 2001-290487/30.

XX Improving the efficacy of treatments involving the administration of

PT interferon-alpha by co-administering an isolated immunostimulatory

PT nucleic acid -

XX Example 17; Page 167; 168pp; English.

XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.

XX Sequence 20 BP; 1 A; 3 C; 13 G; 3 T; 0 other;

Query Match 91.6%; Score 17.4; DB 22; Length 20;
Best Local Similarity 94.7%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX Hartmann G, Bratzler RL, Krieg A;
 XX WPI; 2001-290487/30.
 DR
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 PT
 XX Claim 201; Page 103; 168pp; English.
 PS
 XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX
 SQ Sequence 19 BP; 2 A; 3 C; 12 G; 2 T; 0 other;
 Query Match 100.0%; Score 19; DB 22; Length 19;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ggggacgtcgactggggg 19
 Db 1 ggggacgtcgactggggg 19
 RESULT 2
 AAF99843
 ID AAF99843 standard; DNA; 19 BP.
 XX AAF99843;
 AC
 XX 12-JUN-2001 (first entry)
 DT
 XX Immunostimulatory nucleic acid #959.
 DE
 XX Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX
 OS Synthetic.
 XX
 WO200122972-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 25-SEP-2000; 2000WO-US26383.
 PF
 XX 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX
 XX (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 PA
 XX Krieg AM, Schetter C, Vollmer J;
 XX WPI; 2001-273485/28.
 DR
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 PT
 XX Claim 101; Page 59; 338pp; English.
 PS

XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious diseases, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 XX
 SQ Sequence 19 BP; 2 A; 3 C; 12 G; 2 T; 0 other;
 Query Match 100.0%; Score 19; DB 22; Length 19;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ggggacgtcgactggggg 19
 Db 1 ggggacgtcgactggggg 19
 RESULT 3
 AAF98871
 ID AAF98871 standard; DNA; 20 BP.
 XX AAF98871;
 AC
 XX 11-JUN-2001 (first entry)
 DT
 XX Immunostimulatory nucleic acid assay control oligo SEQ ID NO: 152.
 DE
 XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
 KW viral infection; phosphorothioate backbone; palindromic; cancer; ds.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH modified_base 1..20
 FT /*tag= a
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 XX
 WO200122990-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 27-SEP-2000; 2000WO-US26527.
 PF
 XX 27-SEP-1999; 99US-0156147.
 PR
 XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 PA
 XX Hartmann G, Bratzler RL, Krieg A;
 PI
 XX WPI; 2001-290487/30.
 DR
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 PT
 XX Example 17; Page 163; 168pp; English.
 PS
 XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:21:53 ; Search time 1145.36 seconds
(without alignments)
28.481 Million cell updates/sec

Title: US-09-672-126-30

Perfect score: 19

Sequence: 1 gggagctgcagctggggg 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	19	100.0	19	22 AAF98760 Human IFN-alpha im
2	19	100.0	19	22 AAF98760 Immunostimulatory
3	19	100.0	20	22 AAF98760 Immunostimulatory
4	19	100.0	20	22 AAF98760 Immunostimulatory
5	17.4	91.6	20	22 AAF98880 Immunostimulatory
6	17.4	91.6	20	22 AAF98870 Immunostimulatory
7	17.4	91.6	1170	13 AAQ20217 Sequence of tu3 g
8	16	84.2	701	21 AAF12250 Aspergillus oryzae
9	15.8	83.2	20	22 AAF98748 Human IFN-alpha im

10	15.8	83.2	20	22 AAF99768 Immunostimulatory
11	15.8	83.2	20	22 AAF99830 Immunostimulatory
12	15.8	83.2	21	22 AAF98767 Human IFN-alpha im
13	15.8	83.2	21	22 AAF99873 Immunostimulatory
14	15.8	83.2	228	23 ABL25165 Drosophila melanog
15	15.8	83.2	1974	22 ABL09482 Human secreted pro
16	15.8	83.2	2082	21 ABL69802 Human breast tumou
17	15.8	83.2	2288	23 ABL25164 Drosophila melanog
18	15.8	83.2	3384	22 AAS59854 Human novel cytol
19	15.8	83.2	3384	23 AAS77630 DNA encoding novel
20	15.8	83.2	3384	23 AAS88731 DNA encoding novel
21	15.8	83.2	4416	23 AAS86058 DNA encoding novel
22	15.8	83.2	37716	23 AAS59553 Propionibacterium
23	15.4	81.1	628	22 AAL02444 Human reproductive
24	15.4	81.1	900	21 AAC77507 Human ORFX ORF3062
25	15.4	81.1	9507	22 AAL07097 Human reproductive
26	15.4	81.1	44377	18 AAT78508 Platenolide syntha
27	15.4	81.1	44377	18 AAT80414 Human IFN-alpha im
28	14.8	77.9	20	22 AAF98763 Immunostimulatory
29	14.8	77.9	20	22 AAF98879 Immunostimulatory
30	14.8	77.9	20	22 AAF99866 Immunostimulatory
31	14.8	77.9	20	22 AAF99868 Probe used to dete
32	14.8	77.9	50	22 AAH44830 Human secreted pro
33	14.8	77.9	218	21 AAC19351 Novel human diagno
34	14.8	77.9	278	22 AAS39076 Human polynucleoti
35	14.8	77.9	309	22 AAI80217 Human brain Expres
36	14.8	77.9	333	14 AAQ61222 Human ORFX ORF767
37	14.8	77.9	364	21 AAC75212 Novel human polynu
38	14.8	77.9	429	22 AAF66595 Zsea mays DNA fragm
39	14.8	77.9	459	21 AAC39699 Human reproductive
40	14.8	77.9	464	22 AAL01055 Human reproductive
41	14.8	77.9	486	22 AAL04681 Drosophila melanog
42	14.8	77.9	633	23 ABL06207 DNA encoding novel
43	14.8	77.9	750	23 AAS84700 Human immune/haema
44	14.8	77.9	853	22 AAK79335 Human polynucleoti
45	14.8	77.9	1040	22 AA161179

ALIGNMENTS

RESULT 1

ID	AAF98760	standard; DNA; 19 BP.
XX	AAF98760;	
AC	AAF98760;	
XX	11-JUN-2001	(first entry)
DT	Human IFN-alpha immunostimulatory nucleic acid	SEQ ID NO: 30.
DE	Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;	
KW	viral infection; phosphorothioate backbone; palindrome; cancer; ds.	
XX	Synthetic.	
OS	Key	Location/Qualifiers
XX	modified_base	1..2
FT	/*tag= a	
FT	/mod_base= "OTHER"	
FT	/note= "phosphorothioate linkage"	
FT	modified_base	15..18
FT	/*tag= b	
FT	/mod_base= "OTHER"	
FT	/note= "phosphorothioate linkage"	
XX	WO200122990-A2.	
PN	05-APR-2001.	
PD	27-SEP-2000; 2000WO-US26527.	
XX	27-SEP-1999; 99US-0156147.	
PR		

* 163655 192169: contig of 28515 bp in length.

FEATURES
source
1..192169
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP21-43909"
/clone_lib="RPCI mouse PAC library 21"
BASE COUNT 52163 a 42050 c 41863 g 50666 t 5427 others
ORIGIN

Query Match 86.3%; Score 16.4; DB 2; Length 192169;
Best Local Similarity 94.4%; Pred. No. 1.4e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gggacgtcgacgtggggg 19
||||| |||||||
Db 166637 GGGACGTCGCGGGG 166620

RESULT 14
AX104781
LOCUS AX104781 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 973 from Patent WO0122972.
ACCESSION AX104781
VERSION AX104781.1 GI:13920978
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg.A.M., Schetter.C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 973 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630" 2 t
BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match 83.2%; Score 15.8; DB 6; Length 20;
Best Local Similarity 89.5%; Pred. No. 1.4e+04;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggacgtcgacgtggggg 19
||||| |||||||
Db 1 GGGTCGTCGACGAGGGG 19

RESULT 15
AX104844
LOCUS AX104844 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1036 from Patent WO0122972.
ACCESSION AX104844
VERSION AX104844.1 GI:13921041
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg.A.M., Schetter.C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 1036 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"

```

ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheraia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE          1 (bases 1 to 192169)
JOURNAL        DOE Joint Genome Institute.
AUTHORS        Sequencing of Mouse
TITLE          Unpublished
JOURNAL        2 (bases 1 to 192169)
AUTHORS        DOE Joint Genome Institute.
TITLE          Direct Submission
JOURNAL        Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint
COMMENT        Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
                On Jul 15, 2000 this sequence version replaced gi:6980195.
                -----Genome Center Sequencing
                Center: Joint Genome Institute
                Center Code: JGI
                Web site: http://www.jgi.doe.gov
                -----
                Project Information
                Center Project Name: 1426557
                Center clone name: RPCI-21_43909
                -----
                Summary Statistics
                Consensus quality: 148566 bases at least Q40
                Consensus quality: 170600 bases at least Q30
                Consensus quality: 175179 bases at least Q20
                Estimated insert size: 180000; pulse field gel estimation
                Estimated insert size: 186769; sum-of-contigs estimation
                Quality coverage: 3.49 in Q20 bases; pulse field gel estimation
                Quality coverage: 3.37 in Q20 bases; sum-of-contigs estimation.
                NOTE: This is a 'working draft' sequence. It currently
                * consists of 55 contigs. The true order of the pieces
                * is not known and their order in this sequence record is
                * arbitrary. Gaps between the contigs are represented as
                * runs of N, but the exact sizes of the gaps are unknown.
                * This record will be updated with the finished sequence.
                * as soon as it is available and the accession number will
                * be preserved.
                *
                * 1
                * 1236: contig of 1236 bp in length
                * 1337: gap of unknown length
                * 1337: contig of 1374 bp in length
                * 2811: gap of unknown length
                * 2811: contig of 1204 bp in length
                * 4015: gap of unknown length
                * 4015: contig of 1290 bp in length
                * 5405: gap of unknown length
                * 5405: contig of 1752 bp in length
                * 7257: gap of unknown length
                * 7257: contig of 1473 bp in length
                * 8830: gap of unknown length
                * 8830: contig of 1312 bp in length
                * 10242: gap of unknown length
                * 10242: contig of 1671 bp in length
                * 12013: gap of unknown length
                * 12013: contig of 1494 bp in length
                * 13607: gap of unknown length
                * 13607: contig of 1346 bp in length
                * 15053: gap of unknown length
                * 15053: contig of 1005 bp in length
                * 16158: gap of unknown length
                * 16158: contig of 1323 bp in length
                * 17581: gap of unknown length
                * 17581: contig of 1320 bp in length
                * 19001: gap of unknown length
                * 19001: contig of 1852 bp in length
                * 20953: gap of unknown length
                * 20953: contig of 1814 bp in length
                * 22866: gap of unknown length
                * 22866: contig of 2304 bp in length
                * 22967: gap of unknown length
                * 22967: contig of 1437 bp in length
                * 25371: gap of unknown length
                * 25371: contig of 1440 bp in length
                * 26908: contig of 1440 bp in length
                *
                * 28348: gap of unknown length
                * 28448: contig of 2011 bp in length
                * 30558: gap of unknown length
                * 31686: contig of 1128 bp in length
                * 31687: gap of unknown length
                * 31787: gap of unknown length
                * 33239: contig of 1452 bp in length
                * 33339: gap of unknown length
                * 35319: contig of 1980 bp in length
                * 35419: gap of unknown length
                * 36985: contig of 1567 bp in length
                * 37086: gap of unknown length
                * 37086: contig of 2076 bp in length
                * 39162: gap of unknown length
                * 39262: contig of 2443 bp in length
                * 41705: gap of unknown length
                * 41805: contig of 2558 bp in length
                * 44363: gap of unknown length
                * 44463: contig of 1991 bp in length
                * 46453: gap of unknown length
                * 46553: contig of 2392 bp in length
                * 48945: gap of unknown length
                * 48946: contig of 1443 bp in length
                * 50489: gap of unknown length
                * 50589: contig of 1427 bp in length
                * 52015: gap of unknown length
                * 52116: contig of 2005 bp in length
                * 54121: gap of unknown length
                * 54221: contig of 1336 bp in length
                * 55557: gap of unknown length
                * 55557: contig of 2373 bp in length
                * 58030: gap of unknown length
                * 58130: contig of 3395 bp in length
                * 61525: gap of unknown length
                * 61525: contig of 2568 bp in length
                * 64193: gap of unknown length
                * 64293: contig of 1316 bp in length
                * 65608: gap of unknown length
                * 65709: contig of 3041 bp in length
                * 68750: gap of unknown length
                * 68850: contig of 2496 bp in length
                * 71346: gap of unknown length
                * 71446: contig of 4041 bp in length
                * 75487: gap of unknown length
                * 75487: contig of 2921 bp in length
                * 78508: gap of unknown length
                * 78608: contig of 4200 bp in length
                * 82808: gap of unknown length
                * 82908: contig of 3831 bp in length
                * 86739: gap of unknown length
                * 86839: contig of 4196 bp in length
                * 91035: gap of unknown length
                * 91135: contig of 4484 bp in length
                * 95619: gap of unknown length
                * 95719: contig of 4500 bp in length
                * 100219: gap of unknown length
                * 100319: contig of 4232 bp in length
                * 104551: gap of unknown length
                * 104551: contig of 4564 bp in length
                * 109215: gap of unknown length
                * 109315: contig of 4665 bp in length
                * 113800: gap of unknown length
                * 114080: contig of 5419 bp in length
                * 119499: gap of unknown length
                * 119599: contig of 6370 bp in length
                * 125969: gap of unknown length
                * 126069: contig of 7772 bp in length
                * 133841: gap of unknown length
                * 133841: contig of 8666 bp in length
                * 142607: gap of unknown length
                * 142707: contig of 8638 bp in length
                * 151345: gap of unknown length
                * 151445: contig of 12110 bp in length
                * 163555: gap of unknown length

```

* 8379 12420: contig of 3942 bp in length
 * 12321 12420: gap of unknown length
 * 12421 19193: contig of 6773 bp in length
 * 19194 19293: gap of unknown length
 * 19294 33318: contig of 14025 bp in length
 * 33319 33418: gap of unknown length
 * 33419 43092: contig of 9674 bp in length
 * 43093 43192: gap of unknown length
 * 43193 61341: contig of 18149 bp in length
 * 61342 61441: gap of unknown length
 * 61442 98529: contig of 37088 bp in length
 * 98530 98629: gap of unknown length
 * 98630 142455: contig of 43826 bp in length
 * 142456 142555: gap of unknown length
 * 142556 208531: contig of 65976 bp in length.

FEATURES

Source

1..208531
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP23-448g13"
 /clone_lib="RP23"

BASE COUNT 61344 a 43196 c 42697 g 60280 t 1024:others
 ORIGIN

Query Match 91.6%; Score 17.4; DB 2; Length 208531;

Best Local Similarity 94.7%; Pred. No. 5.2e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19

|||||

Db 12470 GGGGACGTCGACGAGGGG 12488

RESULT 11

AP000907/c 87802 bp DNA linear PRI 23-MAY-2001
 LOCUS Homo sapiens genomic DNA, chromosome 11q clone:RP11-708L7, complete
 DEFINITION sequences.

ACCESSION AP000907.5 GI:14189747

VERSION

KEYWORDS HTG.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (sites)

1 (sites) Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens genomic DNA

TITLE Published Only in Database (1999) In press

REFERENCE 2 (bases 1 to 87802)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Direct Submission

TITLE Submitted (17-DEC-1999) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan

1-7-22 Sushiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:hattori@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,

Tel:81-45-503-9111, Fax:81-45-503-9170)

On May 22, 2001 this sequence version replaced gi:9757502.

Location/Qualifiers

1..87802

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q"

/clone="RP11-708L7"

BASE COUNT 23901 a 19389 c 19536 g 24976 t

ORIGIN

Query Match

86.3%; Score 16.4; DB 9; Length 87802;

Best Local Similarity 94.4%; Pred. No. 1.6e+03;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 18

|||||

Db 79123 GGGGACGTCGACGAGGGG 79106

RESULT 12

AP003615/c 180397 bp DNA linear HTG 11-MAY-2001
 LOCUS Oryza sativa chromosome 6 clone P0486H12, *** SEQUENCING IN
 DEFINITION PROGRESS ***, in ordered pieces.

ACCESSION AP003615.1 GI:14020953

VERSION

KEYWORDS HTG; HTGS_PHASE2.

SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:P0486H12.

ORGANISM Oryza sativa

REFERENCE 1 (sites)

1 (sites) Sasaki,T., Matsumoto,T. and Yamamoto,K.

Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC

clone:P0486H12

JOURNAL Published Only in Database (2001) In press

REFERENCE 2 (bases 1 to 180397)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Direct Submission

JOURNAL Submitted (10-MAY-2001) Takuji Sasaki, National Institute of

Agrobiological Resources, Rice Genome Research Program; Kannondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail:tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,

Tel:81-298-38-7441, Fax:81-298-38-7468)

NOTE: It currently consists of 1 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces is believed

to be correct as given, however the sizes of the gaps between them

are based on estimates that have provided by the submitter. This

sequence will be replaced by the finished sequence as soon as it is

available and the accession number will be preserved.

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES

Source

1..180397

/organism="Oryza sativa"

/cultivar="Nipponbare"

/db_xref="taxon:4530"

/chromosome="6"

/clone="P0486H12"

BASE COUNT 49502 a 40669 c 40215 g 49861 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 2; Length 180397;

Best Local Similarity 94.4%; Pred. No. 1.4e+03;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 18

|||||

Db 129516 GGGGACGTCGACGTCGGG 129499

RESULT 13

AC020849/c 192169 bp DNA linear HTG 15-JUL-2000
 LOCUS Mus musculus clone RP21-43909, WORKING DRAFT SEQUENCE, 55 unordered
 DEFINITION pieces.

ACCESSION AC020849

VERSION AC020849.4 GI:9211211

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE house mouse.

*	1	8045:	contig of 8046 bp in length
*	8047	1816:	gap of unknown length,
*	8147	12661:	contig of 4515 bp in length
*	12662	12761:	gap of unknown length
*	12762	14003:	contig of 1242 bp in length
*	14004	14103:	gap of unknown length
*	14104	19565:	contig of 5462 bp in length
*	19566	19665:	gap of unknown length
*	19666	25955:	contig of 6291 bp in length
*	25957	26056:	gap of unknown length
*	26057	33462:	contig of 7406 bp in length
*	33463	33562:	gap of unknown length
*	33563	39766:	contig of 6204 bp in length
*	39767	39866:	gap of unknown length

* * * be preserved.

* 1	2320: contig of 2320 bp in length
* 2321	2420: gap of unknown length
* 2421	4909: contig of 2489 bp in length
* 4910	5009: gap of unknown length
* 5010	8278: contig of 3269 bp in length
* 8279	8378: gap of unknown length
* *	

```

Query Match          91.6%; Score 17.4; DB 6; Length 20;
Best Local Similarity 94.7%; Pred. No. 3e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
|||||
Db 1 GGGGACGTGCTGCTGGGG 19

RESULT 6
LOCUS AX105262 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 161 from Patent WO0122990.
ACCESSION AX105262
VERSION AX105262.1 GI:13921412
KEYWORDS synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 161 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
misc_feature 1..2
/note="Backbone has phosphorothioate linkages."
misc_feature 3..14
/note="Backbone has phosphodiester linkages."
misc_feature 15..19
/note="Backbone has phosphorothioate linkages."
misc_feature 20
/note="Backbone has phosphodiester linkages."
BASE COUNT 1 a 3 c 13 g 3 t
ORIGIN

Query Match          91.6%; Score 17.4; DB 6; Length 20;
Best Local Similarity 94.7%; Pred. No. 3e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
|||||
Db 1 GGGGACGTGCTGCTGGGG 19

RESULT 7
LOCUS SRTUF3 2731 bp DNA linear BCT 09-JAN-1995
DEFINITION S.ramocissimus tuf3 gene for elongation factor Tu3.
ACCESSION X67059
VERSION X67059.1 GI:47487
KEYWORDS elongation factor; elongation factor Tu3; tuf3 gene.
SOURCE Streptomyces ramocissimus.
ORGANISM Streptomyces ramocissimus
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
1 (bases 1 to 2731)
REFERENCE 1
AUTHORS Vijgenboom,E.
TITLE Direct Submission
JOURNAL Submitted (23-JUN-1992) E. Vijgenboom, John Innes Institute, Colney
Lane, Norwich NR4 7UH, UK
2 (bases 1 to 2731)
REFERENCE 1
AUTHORS Vijgenboom,E., Woudt,L.P., Heinstra,P.W.H., Rietveld,K., van
Haarlem,J., van Wezel,G.P., Shochat,S. and Bosch,L.
TITLE Three tuf-like genes in the kirromycin producer Streptomyces

```

```

ramocissimus
Microbiology 40, 983-998 (1994)
Location/Qualifiers
1..2731
/organism="Streptomyces ramocissimus"
/db_xref="taxon:1925"
1111..1114
1122..2291
/gene="tuf3"
1122..2291
/gene="tuf3"
/codon_start=1
/transl_table=11
/product="elongation factor Tu3"
/protein_id="CAA4744.1"
/db_xref="GI:47488"
/db_xref="SWISS-PROT:P29544"
/translation="MSKTAYVVRTPHLNIGTMGHVDHGKTTLTAAITKVLAEKSGTFF
VPDRIDRAPEEAARGITINIAHVEYEDTRHYAHVDMPGHADYVKNVVTGAOLDGA
ILVVSALDGIIMPOTAHEVLLAROVGDHIVVALNKADAGDELDLVELEVRDLSEH
GYGGGAPVVRVSGIKALEGDPKWTASIEALLDAVDYVPMPEKYVAPFLPVENVL
TITGRGTIVTGAVERGVTVGNRVEVLGAGLETVTGLETEGKPMDEQAQGNVALLL
RGVPRDVRRHGHVVAAPGVSVPFRSAQVYLSAREGGRTPTVTSYGPQFYIRTD
VVGVDVLDGEVGVARPGETVSMIVELGREVPLEPGLGFAIREGRTVGAGTVALV"
BASE COUNT 410 a 934 c 990 g 397 t
ORIGIN

Query Match          91.6%; Score 17.4; DB 1; Length 2731;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
|||||
Db 2130 GGGGACGTGCGACCTGGGG 2148

RESULT 8
LOCUS AC108753 131278 bp DNA linear HTG 31-JAN-2002
DEFINITION Oryza sativa chromosome 9 clone OSJNBa0010B06, *** SEQUENCING IN
PROGRESS ***, 5 unordered pieces.
ACCESSION AC108753.1 GI:18449959
VERSION AC108753
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 131278)
REFERENCE 1
AUTHORS Yun,D.-W., Hahn,J.-H., Yoon,U.-H., Lee,J.-S., Lee,M.-C., Eun,M.Y.
and Kim,H.-I.
TITLE Oryza sativa BAC OSJNBa0010B06 genomic sequence
JOURNAL Unpublished
2 (bases 1 to 131278)
AUTHORS Hahn,J.-H. and Kim,H.-I.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2002) Rice Genome Sequencing Project, National
Institute of Agricultural Science and Technology(NIAST), RDA, 249
Seodun-dong, Suwon 441-707, Korea (E-mail:jhhahn@rda.go.kr,
Tel:82-31-290-0309, Fax:82-31-290-0308)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 50005: contig of 50005 bp in length
* 50006 50105: gap of unknown length
* 50106 52284: contig of 2179 bp in length

```

Query Match 100.0%; Score 19; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
Db 1 GGGGACGTCGACGTGGGG 19

RESULT 2
AXI05132 AXI05132 19 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 30 from Patent WO0122990.
DEFINITION AXI05132
ACCESSION AXI05132
VERSION AXI05132.1 GI:13921282
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 19)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced interferon
JOURNAL Patent: WO 0122990-A 30 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)
FEATURES
source Location/Qualifiers
1..19
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 1..2
/note="Backbone has phosphorothioate linkages."
misc_feature 3..14
/note="Backbone has phosphodiester linkages."
misc_feature 15..18
/note="Backbone has phosphorothioate linkages."
misc_feature 19
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 3 c 12 g 2 t
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
Db 1 GGGGACGTCGACGTGGGG 19

RESULT 3
AXI04780 AXI04780 20 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 972 from Patent WO0122972.
DEFINITION AXI04780
ACCESSION AXI04780
VERSION AXI04780.1 GI:13920977
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 972 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical GmbH (DE)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
Db 1 GGGGACGTCGACGTGGGG 19

RESULT 4
AXI05253 AXI05253 20 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 152 from Patent WO0122990.
DEFINITION AXI05253
ACCESSION AXI05253
VERSION AXI05253.1 GI:13921403
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced interferon
JOURNAL Patent: WO 0122990-A 152 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 1..20
/note="Synthetic Oligonucleotide"
BASE COUNT 2 a 3 c 13 g 2 t
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgtcgacgtgggg 19
Db 1 GGGGACGTCGACGTGGGG 19

RESULT 5
AXI04884 AXI04884 20 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 1076 from Patent WO0122972.
DEFINITION AXI04884
ACCESSION AXI04884
VERSION AXI04884.1 GI:13921081
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 1076 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical GmbH (DE)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: August 10, 2002, 02:58:29 ; Search time 2778.35 seconds
(without alignments)
143.108 Million cell updates/sec

Title: US-09-672-126-30
Perfect score: 19
Sequence: 1 ggggacgtcgacgtgggg 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	19	100.0	19	6	AXI04857
2	19	100.0	19	6	AXI05132
3	19	100.0	20	6	AXI04780
4	19	100.0	20	6	AXI05253
5	17.4	91.6	20	6	AXI04884
6	17.4	91.6	20	6	AXI05262
7	17.4	91.6	2731	1	SRTUF3
8	17.4	91.6	131278	2	AC108753
9	17.4	91.6	146436	2	AC108759
10	17.4	91.6	208531	2	AC087560
c 11	16.4	86.3	87802	9	AP000907
c 12	16.4	86.3	180397	2	AP003615
c 13	16.4	86.3	192169	2	AC020849
14	15.8	83.2	20	6	AXI04781
15	15.8	83.2	20	6	AXI04844
16	15.8	83.2	20	6	AXI05120
17	15.8	83.2	21	6	AXI04887
18	15.8	83.2	21	6	AXI05139
19	15.8	83.2	1076	5	GGCMSO1
20	15.8	83.2	1156	14	AF188661
21	15.8	83.2	1156	14	ORVP7
c 22	15.8	83.2	1195	8	AF038326
c 23	15.8	83.2	1431	10	RATPRCG6
24	15.8	83.2	1586	9	AK055889
c 25	15.8	83.2	1705	1	AF031242
c 26	15.8	83.2	1711	1	P26BPO
27	15.8	83.2	2274	9	AK056717
c 28	15.8	83.2	2759	10	AF057702
29	15.8	83.2	3369	9	BC012476
30	15.8	83.2	5413	1	HVU95374
31	15.8	83.2	9540	14	AF448220
32	15.8	83.2	11251	1	AE004635
33	15.8	83.2	12928	1	AE005076
34	15.8	83.2	19636	2	AC109552
c 35	15.8	83.2	30853	2	AC094245
36	15.8	83.2	33779	1	SCGD3
37	15.8	83.2	35710	2	AC103128
c 38	15.8	83.2	36307	9	HS366D1
c 39	15.8	83.2	39726	1	SC8D11
c 40	15.8	83.2	40356	1	SCL6
c 41	15.8	83.2	45370	5	FRU009961
42	15.8	83.2	48881	2	AC094398
43	15.8	83.2	50511	9	AC005214
44	15.8	83.2	52611	2	AC100025
c 45	15.8	83.2	54841	2	AC017535

ALIGNMENTS

RESULT 1	AXI04857	AXI04857	Sequence 1049 from Patent WO0122972.	19 bp	DNA	linear	PAT 30-APR-2001
LOCUS	AXI04857						
DEFINITION	AXI04857						
ACCESSION	AXI04857						
VERSION	AXI04857.1	GI:13921054					
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
FEATURES							
source							
BASE COUNT							
ORIGIN							

AXI04857 Sequence 1049 from Patent WO0122972.
AXI04857
AXI04857.1 GI:13921054
synthetic construct.
synthetic construct.
artificial sequence.
1 (bases 1 to 19)
Krieg, A.M., Schetter, C. and Vollmer, J.C.
Immunostimulatory nucleic acids
Patent: WO 0122972-A 1049 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
Location/Qualifiers
1. 19
/organism="synthetic construct"
/db_xref="taxon:32630" 2 t
2 a 3 c 12 g

; LOCATION: 723...1097
US-08-395-800A-1

Query Match 65.0%; Score 15.6; DB 1; Length 1144;
Best Local Similarity 81.8%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcagctacgtcagggg 22
||||| ||||| |||||
Db 653 GGGTCCACGTGCGCGGGGG 674

RESULT 15
5352575-4/c
; Patent No. 5352575
; APPLICANT: PETROVSKIS, ERIK A.; POST, LEONARD E.; TIMMINS, JAMES G.
; TITLE OF INVENTION: PSEUDORABIES VIRUS PROTEIN
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/513,282
; FILING DATE: 20-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 100,817
; FILING DATE: 29-JUN-1987
; APPLICATION NUMBER: 886,260
; FILING DATE: 16-JUL-1986
; APPLICATION NUMBER: 784,787
; FILING DATE: 04-OCT-1985
; APPLICATION NUMBER: 801,799
; FILING DATE: 26-NOV-1985
; APPLICATION NUMBER: 844,113
; FILING DATE: 26-MAR-1986
; SEQ ID NO: 4
; LENGTH: 1209
5352575-4

Query Match 65.0%; Score 15.6; DB 6; Length 1209;
Best Local Similarity 81.8%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcagctacgtcagggg 22
||||| ||||| |||||
Db 159 GGGCCGACGAGGCGGAGGGG 138

Search completed: August 10, 2002, 03:06:24
Job time: 16050 sec

||||| || ||||| |||||
Db 10041 GGTGGAAGTCGTCGAGGGG 10021

RESULT 12

US-09-102-204-2
; Sequence 2, Application US/09102204
; Patent No. 6150899
; GENERAL INFORMATION:
; APPLICANT: Jones, Brian E.
; APPLICANT: Van Der Kleij, Wilhelmus A.H.
; APPLICANT: Van Solingen, Piet
; APPLICANT: Weyler, Walter
; TITLE OF INVENTION: No. 6190899el Cellulase Producing
; TITLE OF INVENTION: Actinomycetes, Cellulase Produced Therefrom
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,204
; FILING DATE: 22-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/974,041
; FILING DATE: 19-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Stone, Christopher L.
; REGISTRATION NUMBER: 35,696
; REFERENCE/DOCKET NUMBER: GC539
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-846-7555
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1059 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-102-204-2

Query Match 66.7%; Score 16; DB 4; Length 1059;
Best Local Similarity 79.2%; Pred. No. 73;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 999gtcagctgctgagggggg 24
||| || ||||| |||||
Db 324 GGTGTCACCTAGTCGAGGGG 347

RESULT 13

US-09-160-496-4/c
; Sequence 4, Application US/09160496
; Patent No. 6346613
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J
; APPLICANT: Cagney, Gerard
; TITLE OF INVENTION: Composition and Method for Enhancing Paracellular
; TITLE OF INVENTION: Transport across Cell Layers
; FILE REFERENCE: Docket No. 6346613: 98.1070.US
; CURRENT APPLICATION NUMBER: US/09/160,496
; CURRENT FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: US 60/059,644
; EARLIER FILING DATE: 1997-09-24

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Gallus gallus
; PUBLICATION INFORMATION:
; TITLE: Occludin: A novel integral membrane protein localizing
; TITLE: at tight junctions
; JOURNAL: J. Cell Biol.
; VOLUME: 123
; ISSUE: 6
; PAGES: 1777-1788
; DATE: Dec 1993
; DATABASE ACCESSION NUMBER: D21837
US-09-160-496-4

Query Match 66.7%; Score 16; DB 4; Length 1920;
Best Local Similarity 79.2%; Pred. No. 71;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 999gtcagctgctgagggggg 24
||||| ||||| |||||
Db 89 GGGGGCGCGTACCGCTGGGGG 66

RESULT 14

US-08-395-800A-1
; Sequence 1, Application US/08395800A
; Patent No. 5807732
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B
; APPLICANT: LENNON, GREGORY
; APPLICANT: ROQUIER, SYLVIE
; APPLICANT: GIORGI, DOMINIQUE
; APPLICANT: KELLY, ROBERT J
; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE
; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE
; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,800A
; FILING DATE: 28-FEB-1995
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56...721
; FEATURE:
; NAME/KEY: CDS

APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002120
CURRENT APPLICATION NUMBER: US/09/320,878A
CURRENT FILING DATE: 1999-05-27 OF 09/141,908
EARLIER APPLICATION NUMBER: CIP OF 09/073,538
EARLIER FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: CIP OF 09/073,538
EARLIER FILING DATE: 1998-05-06
EARLIER APPLICATION NUMBER: CIP OF 08/846,247
EARLIER FILING DATE: 1997-04-30
EARLIER APPLICATION NUMBER: 60/119,139
EARLIER FILING DATE: 1999-02-08
EARLIER APPLICATION NUMBER: 60/100,880
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: 60/087,080
EARLIER FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 38506
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-320-878-19

Query Match 69.2%; Score 16.6; DB 3; Length 38506;
Best Local Similarity 82.6%; Pred. No. 35;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 gggcgacgtacgtcgagggggg 24
||| ||||| ||| ||||| |
Db 28456 GGGCGACGTCGCGGAGGGGTG 28434

RESULT 10
US-08-125-468-1/c
Sequence 1, Application US/08125468
Patent No. 5589385
GENERAL INFORMATION:
APPLICANT: Ryan, Michael J.
APPLICANT: Lotvin, Jason A.
APPLICANT: Strathy, Nancy
APPLICANT: Fantini, Susan E.
TITLE OF INVENTION: Cloning of the biosynthetic pathway for
TITLE OF INVENTION: chlortetracycline and tetracycline Formation and cosmid
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/125,468
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Tsevdos, Estelle J
REGISTRATION NUMBER: 31,145
REFERENCE/DOCKET NUMBER: 31,255-02
TELEPHONE: (201)831-3241
TELEFAX: (201)831-3305

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30001 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-125-468-1

Query Match 67.5%; Score 16.2; DB 1; Length 30001;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ggtcgacgtacgtcgagggggg 23
||||| || ||||| |||||
Db 10041 GGTGGAAGTCGTGACGGGG 10021

RESULT 11
US-08-474-933-1/c
Sequence 1, Application US/08474933
Patent No. 5866410
GENERAL INFORMATION:
APPLICANT: Ryan, Michael J.
APPLICANT: Lotvin, Jason A.
APPLICANT: Strathy, Nancy
APPLICANT: Fantini, Susan E.
TITLE OF INVENTION: Cloning of the biosynthetic pathway for
TITLE OF INVENTION: chlortetracycline and tetracycline Formation and cosmid
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/474,933
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,468
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Tsevdos, Estelle J
REGISTRATION NUMBER: 31,145
REFERENCE/DOCKET NUMBER: 31,255-02
TELEPHONE: (201)831-3241
TELEFAX: (201)831-3305
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30001 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-474-933-1

Query Match 67.5%; Score 16.2; DB 2; Length 30001;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ggtcgacgtacgtcgagggggg 23

Db 3324 GGGCCGACGTCCGTGGAGGGTG 3302
|||||
|||||

RESULT 6
US-07-945-283-1/c
; Sequence 1, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The EPO and LIT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1815 No. 5352596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,283
; FILING DATE: 19920911
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ribando, Curtis P.
; REGISTRATION NUMBER: 27976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 309-685-4011 ext. 513
; TELEFAX: 309-685-4128
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8438 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudorabies virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 622..6495
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1099, "g")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1267, "t")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1381, "c")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1566, "c")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(7010, "g")
US-07-945-283-1

Query Match 69.2%; Score 16.6; DB 1; Length 8438;
Best Local Similarity 82.6%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ggggtcgcgtacgtcgagggg 23

Db 1315 GGGGGGACGGATGTGACGGGG 1293
|||||
|||||

RESULT 7
US-09-105-537-30/c
; Sequence 30, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30
; LENGTH: 13842
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-30

Query Match 69.2%; Score 16.6; DB 4; Length 13842;
Best Local Similarity 82.6%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 2 ggggtcgcgtacgtcgagggg 24
|||||
|||||

Db 6600 GGGCCGACGTCCGTGGAGGGTG 6578
|||||
|||||

RESULT 8
US-09-105-537-5/c
; Sequence 5, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 36778
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-5

Query Match 69.2%; Score 16.6; DB 4; Length 36778;
Best Local Similarity 82.6%; Pred. No. 35;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 2 ggggtcgcgtacgtcgagggg 24
|||||
|||||

Db 30314 GGGCCGACGTCCGTGGAGGGTG 30292
|||||
|||||

RESULT 9
US-09-320-878-19/c
; Sequence 19, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.

US-08-998-416-335

Query Match 74.2%; Score 17.8; DB 4; Length 820;
Best Local Similarity 90.5%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 gtcgacgtacgtcgagggggg 24
|||||
Db 196 gtcgacgtacgttcgaggggg 216

RESULT 2

US-07-945-283-3
; Sequence 3, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The EP0 and LIT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1815 No. 3552596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,283
; FILING DATE: 19920911
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ribando, Curtis P
; REGISTRATION NUMBER: 27976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 309-685-4011 ext.513
; TELEFAX: 309-685-4128
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1683 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudorabies virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 211..1440
; OTHER INFORMATION: /product= "early protein 0"
US-07-945-283-3

Query Match 69.2%; Score 16.6; DB 1; Length 1683;
Best Local Similarity 82.6%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcgacgtacgtcgagggggg 23
|||||
Db 531 GGGGGCGACGGATGTCGACGGG 553

RESULT 3

5215881-1/c

; Patent No. 5215881
; APPLICANT: CHEUNG, ANDREW K.
; TITLE OF INVENTION: PSEUDORABIES DIAGNOSIS PROBES
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/537,855
; FILING DATE: 13-JUN-1990
; SEQ ID NO:1:
; LENGTH: 1831
5215881-1

Query Match 69.2%; Score 16.6; DB 6; Length 1831;
Best Local Similarity 82.6%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcgacgtacgtcgagggggg 23
|||||
Db 1317 GGGGGCGACGGATGTCGACGGG 1295

RESULT 4

5215881-3/c
; Patent No. 5215881
; APPLICANT: CHEUNG, ANDREW K.
; TITLE OF INVENTION: PSEUDORABIES DIAGNOSIS PROBES
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/537,855
; FILING DATE: 13-JUN-1990
; SEQ ID NO:3:
; LENGTH: 1831
5215881-3

Query Match 69.2%; Score 16.6; DB 6; Length 1831;
Best Local Similarity 82.6%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcgacgtacgtcgagggggg 23
|||||
Db 1317 GGGGGCGACGGATGTCGACGGG 1295

RESULT 5

US-09-105-537-34/c
; Sequence 34, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 4689
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-34

Query Match 69.2%; Score 16.6; DB 4; Length 4689;
Best Local Similarity 82.6%; Pred. No. 38;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ggggtcgacgtacgtcgagggggg 24

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:06:13 ; Search time 277.54 Seconds
(without alignments)
21.241 Million cell updates/sec

Title: US-09-672-126-25

Perfect score: 24

Sequence: 1 ggggtcgactgctgagggggg 24

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.8	74.2	820	4	US-08-998-416-335
2	16.6	69.2	1683	1	US-07-945-283-3
3	16.6	69.2	1831	6	5215881-1
4	16.6	69.2	1831	6	5215881-3
5	16.6	69.2	4689	4	US-09-105-537-34
6	16.6	69.2	8438	1	US-07-945-283-1
7	16.6	69.2	13842	4	US-09-105-537-30
8	16.6	69.2	36778	4	US-09-105-537-5
9	16.6	69.2	38506	3	US-09-320-878-19
10	16.2	67.5	30001	1	US-08-125-468-1
11	16.2	67.5	30001	2	US-08-474-933-1
12	16.6	66.7	1059	4	US-09-102-204-2
13	16.6	66.7	1920	4	US-09-160-496-4
14	15.6	65.0	1144	1	US-08-395-800A-1
15	15.6	65.0	1209	6	5352575-4
16	15.6	65.0	1213	4	US-09-232-468A-7
17	15.6	65.0	1269	4	US-09-151-592-1
18	15.6	65.0	1719	4	US-09-330-740A-9
19	15.6	65.0	1756	2	US-08-465-640-1
20	15.6	65.0	2115	1	US-08-395-800A-7
21	15.6	65.0	2692	1	US-07-932-454A-2
22	15.6	65.0	16885	1	US-08-390-878-16
23	15.6	65.0	4403765	4	US-09-103-840A-2
24	15.6	65.0	4411529	4	US-09-103-840A-1
25	15.4	64.2	2754	2	US-09-028-361A-1
26	15.2	63.3	4403765	4	US-09-103-840A-2
27	15.2	63.3	4411529	4	US-09-103-840A-1

28	15	62.5	1146	1	US-08-482-385A-1	Sequence 1, Appli
29	15	62.5	1559	2	US-08-160-524A-1	Sequence 1, Appli
30	15	62.5	1684	1	US-07-829-016-3	Sequence 3, Appli
31	15	62.5	1684	1	US-08-487-651-3	Sequence 3, Appli
32	15	62.5	1684	2	US-08-487-645A-3	Sequence 3, Appli
33	15	62.5	1886	1	US-08-461-773-15	Sequence 15, Appli
34	15	62.5	1942	3	US-08-627-907A-3	Sequence 3, Appli
35	15	62.5	2335	4	US-09-387-574-9	Sequence 9, Appli
36	15	62.5	2335	4	US-09-668-096-9	Sequence 9, Appli
37	15	62.5	2728	1	US-08-482-385A-5	Sequence 5, Appli
38	15	62.5	2855	2	US-08-776-597A-1	Sequence 1, Appli
39	15	62.5	2855	2	US-08-693-228-1	Sequence 1, Appli
40	15	62.5	4015	4	US-08-810-009-4	Sequence 4, Appli
41	15	62.5	4695	2	US-08-231-193A-57	Sequence 57, Appli
42	15	62.5	4695	2	US-08-486-273A-57	Sequence 57, Appli
43	15	62.5	4695	3	US-08-940-086A-57	Sequence 57, Appli
44	15	62.5	4695	4	US-08-940-035A-57	Sequence 57, Appli
45	15	62.5	5496	4	US-09-462-284-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-998-416-335
; Sequence 335, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtie, Philipp
; APPLICANT: Rebschue, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 335:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 820 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1265UP

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

1 (bases 1 to 660)

Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation - Center for Medicago Genomics Research

Unpublished (2000)

Contact: Dixon RA

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7302

Fax: 580 221 7380

Email: radixon@noble.org

Insert Length: 660 Std Error: 0.00

Plate: 036 row: A column: 09

Seq primer: TCACACAGGAACACGCTATGAC.

FEATURES

source

1..660
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF036A09BC"

/clone_lib="Elicited cell culture"

/tissue_type="Cell suspensions derived from root tissues"

/dev_stage="Cells were subcultured every 14 days. Cells were induced six days after subculture"

/note="Vector: Lambda zap; Cells were induced with yeast cell wall extracts equivalent to 50ug/ml glucose in the final concentration. Samples were taken at 0.5, 1, 12 and 24 hours after induction. Equal amounts of RNA from each time point were pooled and used for mRNA isolation."

211 a 128 c 171 g 148 t 2 others

BASE COUNT
ORIGIN

Query Match 75.8%; Score 18.2; DB 10; Length 660;

Best Local Similarity 87.0%; Pred. No. 1.6e+03;

Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ggggtcgacgtacgtcgaggggg 24

||||| ||| ||| |||

Db 60 GGGTCGACGGCGGCGAGGGG 38

RESULT 15

AG131484/c

LOCUS

AG131484 Pan troglodytes DNA, clone: PTB-143J02.R, genomic survey sequence. 689 bp DNA linear GSS 04-NOV-2001

ACCESSION AG131484

VERSION AG131484.1 GI:16661162

KEYWORDS GSS; GSS (genome survey sequence).

SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male BAC Library clone:PTB-143J02.R.

ORGANISM

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

1 (sites)

Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,

Totoki,Y., Watanabe,H. and Sakaki,Y.

BAC end sequences of Library PTB

Unpublished

2 (bases 1 to 689)

Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,

Totoki,Y., Watanabe,H. and Sakaki,Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:chimpbes@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,

tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS

Sequencing: M13Rev

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI

FEATURES

source

1..689

/organism="Pan troglodytes"

/db_xref="taxon:9598"

/clone="PTB-143J02.R"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

241 a 165 c 124 g 158 t 1 others

BASE COUNT

ORIGIN

Query Match 75.8%; Score 18.2; DB 12; Length 689;

Best Local Similarity 87.0%; Pred. No. 1.6e+03;

Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggggtcgacgtacgtcgaggggg 23

||||| ||| ||| ||| |||

Db 670 GGGTGGCGGACGACGGCGAGGGG 648

Search completed: August 10, 2002, 02:11:20

Job time: 13141 sec

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaeae; Oryza.

REFERENCE
AUTHORS 1 (bases 1 to 300)
TITLE Large-scale Sequencing Analysis of ESTs from Rice Seedling
JOURNAL Unpublished (1999)
COMMENT Contact: Eun M.Y.

Department of CytoGenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.osti.re.kr.

Location/Qualifiers
1. 300
/organism="Oryza sativa"
/cultivar="Milyang23"
/db_xref="taxon:4530"
/clone="99AS345"

/clone_lib="Rice Seedling Lambda ZAPII cDNA Library"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
/note="Vector: pBluescript SK(+); Site.1: EcoRI; Site.2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site"

BASE COUNT 68 a 62 c 96 g 74 t
ORIGIN

Query Match 75.8%; Score 18.2; DB 9; Length 300;
Best Local Similarity 87.0%; Pred. No. 1.4e+03;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggggtcagctacgtcgagggg 23
||||| ||| ||||| |||||
DB 10 GGGTCGTCGTCGTCGAGGGG 32

RESULT 5
BE316884/c
LOCUS NF056F07LF1059 Developing leaf Medicago truncatula cDNA clone
DEFINITION NF056F07LF 5', mRNA sequence.

ACCESSION BE316884
VERSION BE316884.2 GI:11962453
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 318)
AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula leaf library
Unpublished (2000)
On Jul 14, 2000 this sequence version replaced gi:9190661.
Contact: May GD

Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org

Medicago Genome Initiative accession: MGI:S:21085
Insert Length: 770 Std Error: 0.00
Plate: 056 row: F column: 07
Seq primer: TCACAGGAAACAGCATGAC.

Location/Qualifiers
1. 300
/organism="Medicago truncatula"
/cultivar="Jemalong"
/db_xref="taxon:3880"
/clone="MtBA08F02"
/clone_lib="MtBA"

/tissue_type="root tips"
/dev_stage="harvested after 3 days of N-starvation"
/note="Vector: pBluescript pSK; Site.1: EcoRI; Site.2:
XhoI; Plants were grown in an aeroponic chamber for 14
days on nitrogen-rich medium followed by 3 days on N-free
medium. RNA was extracted from root tips (1-3 cm). cDNA
was prepared from polyA+ enriched RNA. The cDNA was
directionally ligated into Uni-zapXR vector from
Stratagene and packaged using Gigapack Gold packaging
extracts. Plasmids containing cDNA inserts were
mass-excised from phage stocks using ExSsIII helper phage
and propagated in SOLR cells. Clone ordering and
sequencing was performed by the Centre National de
Sequencage (Genoscope, Evry, France)."

BASE COUNT 105 a 75 c 71 g 82 t
ORIGIN

Query Match 75.8%; Score 18.2; DB 10; Length 318;
Best Local Similarity 87.0%; Pred. No. 1.4e+03;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ggggtcagctacgtcgagggg 24
||||| ||| ||||| |||||
DB 46 GGGTCGACGCGCGAGGGG 24

RESULT 6
AL3366535/c
LOCUS MTBA08F02F1 MtBA Medicago truncatula cDNA clone MTBA08F02 T3, mRNA
DEFINITION AL3366535 sequence.

ACCESSION AL3366535
VERSION AL3366535.1 GI:9666288
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 333)
AUTHORS Journet, E.P., Crespeau, H., van-Tuinen, D., Gouzy, J., Jaillon, O., Niebel, A., Carreau, V., Chatagnier, O., Kahn, D., Gianinazzi-Pearson, V. and Gamas, P.

Medicago truncatula ESTs from nitrogen-starved roots
Unpublished (2000)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Contact : Pascal Gamas and Etienne-Pascal Journet, Laboratoire de
Biologie Moleculaire des Relations Plantes-Microorganismes,
CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :
Mt-est@toulouse.inra.fr Website :
http://sequence.toulouse.inra.fr/Mtruncatula.html).

Location/Qualifiers
1. 333
/organism="Medicago truncatula"
/cultivar="Jemalong"
/db_xref="taxon:3880"
/clone="MtBA08F02"
/clone_lib="MtBA"

/tissue_type="root tips"
/dev_stage="harvested after 3 days of N-starvation"
/note="Vector: pBluescript pSK; Site.1: EcoRI; Site.2:
XhoI; Plants were grown in an aeroponic chamber for 14
days on nitrogen-rich medium followed by 3 days on N-free
medium. RNA was extracted from root tips (1-3 cm). cDNA
was prepared from polyA+ enriched RNA. The cDNA was
directionally ligated into Uni-zapXR vector from
Stratagene and packaged using Gigapack Gold packaging
extracts. Plasmids containing cDNA inserts were
mass-excised from phage stocks using ExSsIII helper phage
and propagated in SOLR cells. Clone ordering and
sequencing was performed by the Centre National de
Sequencage (Genoscope, Evry, France)."

source

1. 318

/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF056F07LF"
/clone_lib="Developing leaf"
/tissue_type="leaf"

/dev_stage="Pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of very
young, developing, mature and senescing leaves."

BASE COUNT 117 a 65 c 62 g 74 t
ORIGIN

Query Match 75.8%; Score 18.2; DB 10; Length 318;
Best Local Similarity 87.0%; Pred. No. 1.4e+03;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ggggtcagctacgtcgagggg 24
||||| ||| ||||| |||||
DB 46 GGGTCGACGCGCGAGGGG 24

RESULT 6
AL3366535/c
LOCUS MTBA08F02F1 MtBA Medicago truncatula cDNA clone MTBA08F02 T3, mRNA
DEFINITION AL3366535 sequence.

ACCESSION AL3366535
VERSION AL3366535.1 GI:9666288
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 333)
AUTHORS Journet, E.P., Crespeau, H., van-Tuinen, D., Gouzy, J., Jaillon, O., Niebel, A., Carreau, V., Chatagnier, O., Kahn, D., Gianinazzi-Pearson, V. and Gamas, P.

Medicago truncatula ESTs from nitrogen-starved roots
Unpublished (2000)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Contact : Pascal Gamas and Etienne-Pascal Journet, Laboratoire de
Biologie Moleculaire des Relations Plantes-Microorganismes,
CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :
Mt-est@toulouse.inra.fr Website :
http://sequence.toulouse.inra.fr/Mtruncatula.html).

Location/Qualifiers
1. 333
/organism="Medicago truncatula"
/cultivar="Jemalong"
/db_xref="taxon:3880"
/clone="MtBA08F02"
/clone_lib="MtBA"

/tissue_type="root tips"
/dev_stage="harvested after 3 days of N-starvation"
/note="Vector: pBluescript pSK; Site.1: EcoRI; Site.2:
XhoI; Plants were grown in an aeroponic chamber for 14
days on nitrogen-rich medium followed by 3 days on N-free
medium. RNA was extracted from root tips (1-3 cm). cDNA
was prepared from polyA+ enriched RNA. The cDNA was
directionally ligated into Uni-zapXR vector from
Stratagene and packaged using Gigapack Gold packaging
extracts. Plasmids containing cDNA inserts were
mass-excised from phage stocks using ExSsIII helper phage
and propagated in SOLR cells. Clone ordering and
sequencing was performed by the Centre National de
Sequencage (Genoscope, Evry, France)."

BASE COUNT 105 a 75 c 71 g 82 t
ORIGIN

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; The library was constructed by Dan Howe, University of Kentucky. cDNAs were synthesized from poly(A)+ RNA by oligo d(T) priming and directionally cloned into the Uni-ZAP XR lambda vector. The library was mass excised as phagemids and rescued in SOLR cells. The plasmid library was recovered from the SOLR cells and transformed in mass into DH10B cells for sequencing. WARNING: the library contains a small percentage of cDNAs derived from the bovine host cells."

BASE COUNT 94 a 97 c 83 g 75 t

ORIGIN

Query Match 78.3%; Score 18.8; DB 10; Length 349;

Best Local Similarity 90.9%; Pred. No. 8.4e+02;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ggtcgacgtacgtcgagggggg 24
|||||

Db 315 GGTGACGTACGTCCACGGGGG 294

RESULT 2

LOCUS BE636098/c

DEFINITION BE636098 479 bp mRNA linear EST 25-AUG-2000

ACCESSION SNEST4a15c10.y1 csn 1 S neuropa invitro merozoite cDNA Sarcocystis

VERSION neuropa cDNA 5', mRNA sequence.

KEYWORDS BE636098.1 GI:9918785

SOURCE EST.

ORGANISM Sarcocystis neuropa.

REFERENCE 1 (bases 1 to 479)

AUTHORS Howe,D.K., Stamper,S., Tang,K., Sibley,L.D., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J., Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., McCann,R., Blistain,A., Bennet,J., Schmitt,A., Ronko,I., Tsagarisvilli,R., Fedele,M., Belaygorod,L., Franklin,C., Carr,L.M., Grow,A., Maguire,L., Wadkins,J., Richey,J., Waterston,R. and Wilson,R.

TITLE Sarcocystis neuropa EST project

JOURNAL Unpublished (2000)

COMMENT Contact: Daniel K. Howe
Sarcocystis neuropa EST project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Contact Daniel K. Howe (dkhowe2@pop.uky.edu) for further information relating to organism, libraries, or clone availability.
Seq primer: -40RP from Gibco
High quality sequence stop: 398.

FEATURES

source

1. .479

/organism="Sarcocystis neuropa"

/strain="Sn3"

/db_xref="taxon:42890"

/clone_lib="csn 1 S neuropa invitro merozoite cDNA"

/dev_stage="merozoite"

/lab_host="DH10B"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; The library was constructed by Dan Howe, University of Kentucky. cDNAs were synthesized from poly(A)+ RNA by oligo d(T) priming and directionally cloned into the Uni-ZAP XR lambda vector. The library was mass excised as phagemids and rescued in SOLR cells. The plasmid library was recovered from the SOLR cells and transformed in mass into DH10B cells for sequencing. WARNING: the library contains a small percentage of cDNAs derived from the bovine host cells."

BASE COUNT 126 a 138 c 116 g 98 t 1 others

ORIGIN

Query Match 78.3%; Score 18.8; DB 10; Length 479;

Best Local Similarity 90.9%; Pred. No. 8.8e+02;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ggtcgacgtacgtcgagggggg 24
|||||

Db 308 GGTGACGTACGTCCACGGGGG 287

RESULT 3

LOCUS CNS01MLK/c

DEFINITION CNS01MLK 813 bp DNA linear GSS 14-JUN-2001

ACCESSION from strain PEST of Anopheles gambiae (African malaria mosquito), genomic survey sequence.

VERSION AL151081

KEYWORDS AL151081.1 GI:7011560

SOURCE GSS.

ORGANISM African malaria mosquito.

REFERENCE 1 (bases 1 to 813)

AUTHORS Anopheles gambiae

TITLE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.

JOURNAL Genoscope.

REFERENCE 2 (bases 1 to 813)

AUTHORS Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J.

TITLE Direct Submission

JOURNAL Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr. Roux, Paris 75015, France

COMMENT This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.

FEATURES

Location/Qualifiers

1. .813

/organism="Anopheles gambiae"

/strain="PEST"

/db_xref="taxon:7165"

/clone_lib="21P17"

/clone_lib="NotreDame1"

/note="end : T7"

BASE COUNT 235 a 165 c 166 g 235 t 12 others

ORIGIN

Query Match 78.3%; Score 18.8; DB 12; Length 813;

Best Local Similarity 90.9%; Pred. No. 9.5e+02;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ggtcgacgtacgtcgagggggg 24
|||||

Db 490 GGTGACGTACGTCTGGGGGG 469

RESULT 4

LOCUS BE230113

DEFINITION BE230113 300 bp mRNA linear EST 07-JUL-2000

ACCESSION 99AS345 Rice Seedling Lambda ZAPII cDNA Library Oryza sativa cDNA clone 99AS345, mRNA sequence.

VERSION BE230113

KEYWORDS BE230113.1 GI:8956310

SOURCE EST.

ORGANISM Oryza sativa.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:11:17 ; Search time 9068.22 seconds
(without alignments)
35.721 Million cell updates/sec

Title: US-09-672-126-25

Perfect score: 24

Sequence: 1 ggggtgacgtacgtcgagggggg 24

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_estl.*
10: gb_est2.*
11: gb_htc.*
12: gb_gss.*
13: em_gss_hum.*
14: em_gss_inv.*
15: em_gss_pln.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.8	78.3	349	10 BE636147	BE636147 SnEST4a15
C 2	18.8	78.3	479	10 BE636098	BE636098 SnEST4a15
C 3	18.8	78.3	813	12 CNS01MLK	AL151081 Anopheles
C 4	18.2	75.8	300	9 BE230113	BE230113 99AS345 R
C 5	18.2	75.8	318	10 BE316884	BE316884 NF056F07L
C 6	18.2	75.8	333	9 AL366535	AL366535 MtBA08F02
C 7	18.2	75.8	546	10 BF651137	BF651137 NF101G04E
C 8	18.2	75.8	565	9 AW698853	AW698853 NF109G02S
C 9	18.2	75.8	566	10 BE998033	BE998033 EST429756
C 10	18.2	75.8	580	9 AW585247	AW585247 N211530e
C 11	18.2	75.8	586	10 BF645320	BF645320 NF037A05E
C 12	18.2	75.8	601	10 BE998032	BE998032 EST429755
C 13	18.2	75.8	628	10 BF636231	BF636231 NF106H06D
C 14	18.2	75.8	660	10 BF645238	BF645238 NF036A09E
C 15	18.2	75.8	689	12 AG131484	AG131484 Pan trogl
C 16	18.2	75.8	1203	10 BG819508	BG819508 602783331
C 17	17.6	73.3	224	9 AV345422	AV345422 AV345422

18	17.6	73.3	247	10	BG604583
C 19	17.6	73.3	546	12	AQ689909
C 20	17.6	73.3	556	12	AQ854465
C 21	17.6	73.3	778	10	BF830662
C 22	17.6	73.3	863	12	AZ194898
C 23	17.4	72.5	484	12	CNS03SUR
C 24	17.2	71.7	276	10	BG966439
C 25	17.2	71.7	592	10	BF263216
C 26	17.2	71.7	700	9	AL508415
C 27	17.2	71.7	755	12	CNS0416C
C 28	17.2	71.7	880	12	CNS03TQG
C 29	16.8	70.0	519	9	AJ273517
C 30	16.8	70.0	592	10	BG604955
C 31	16.8	70.0	732	12	AZ573320
C 32	16.8	70.0	888	12	CNS0272L
C 33	16.8	70.0	925	10	BF028418
C 34	16.8	70.0	926	12	CNS03X70
C 35	16.8	70.0	960	12	AG181755
C 36	16.8	70.0	968	12	CNS03T84
C 37	16.8	70.0	997	12	CNS0213U
C 38	16.8	70.0	1063	12	CNS03B8L
C 39	16.6	69.2	131	9	AW062657
C 40	16.6	69.2	161	9	AV133987
C 41	16.6	69.2	232	10	BI726817
C 42	16.6	69.2	271	9	BB552395
C 43	16.6	69.2	320	9	AA888750
C 44	16.6	69.2	349	10	C26056
C 45	16.6	69.2	353	9	AU082133

ALIGNMENTS

RESULT 1

BE636147/c

LOCUS

DEFINITION

neuroana cDNA 5', mRNA sequence.

ACCESSION

BE636147

VERSION

BE636147.1

KEYWORDS

EST.

SOURCE

Sarcocystis neuroana.

ORGANISM

Eukaryota; Alveolata; Apicomplexa; Coccidia; Elmeriida;

Sarcocystidae; Sarcocystis.

REFERENCE

AUTHORS

TITLE

Sarcocystis neuroana EST project

JOURNAL

Unpublished (2000)

COMMENT

Contact: Daniel K. Howe

Sarcocystis neuroana EST project

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Contact Daniel K. Howe (dkhowe2@pop.uky.edu) for further

information relating to organism, libraries, or clone availability.

Seq primer: -40RP from Gibco

High quality sequence stop: 328.

Location/Qualifiers

1..349

/organism="Sarcocystis neuroana"

/strain="Sn3"

/db_xref="taxon:42890"

/clone_lib="cSn 1 S neuroana invitro merozoite cDNA"

/dev_stage="merozoite"

/lab_host="DH10B"

BE636147 349 bp mRNA linear EST 25-AUG-2000
SnEST4a15h05.y1 csn 1 S neuroana invitro merozoite cDNA Sarcocystis
neuroana cDNA 5', mRNA sequence.

BE636147 GI:9918834

EST.

Sarcocystis neuroana.

Sarcocystis neuroana

Eukaryota; Alveolata; Apicomplexa; Coccidia; Elmeriida;

Sarcocystidae; Sarcocystis.

1 (bases 1 to 349)

Howe, D.K., Stamper, S., Tang, K., Sibley, L.D., Clifton, S., Marra, M.,

Hillier, L., Pape, D., Martin, J., Wyllie, T., Theising, B., Bowers, Y.,

Gibbons, M., Ritter, E., McCann, R., Blistain, A., Bennett, J., Schmitt

, A., Ronko, I., Tsagareishvili, R., Fedele, M., Belaygorod, L.,

Franklin, C., Carr, L.M., Grow, A., Maguire, L., Wadkins, J., Richey, J.,

Waterston, R., and Wilson, R.

Sarcocystis neuroana EST project

Unpublished (2000)

Contact: Daniel K. Howe

Sarcocystis neuroana EST project

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Contact Daniel K. Howe (dkhowe2@pop.uky.edu) for further

information relating to organism, libraries, or clone availability.

Seq primer: -40RP from Gibco

High quality sequence stop: 328.

Location/Qualifiers

1..349

/organism="Sarcocystis neuroana"

/strain="Sn3"

/db_xref="taxon:42890"

/clone_lib="cSn 1 S neuroana invitro merozoite cDNA"

/dev_stage="merozoite"

/lab_host="DH10B"

THIS PAGE BLANK (USPTO)

CC Streptomyces venezuelae ATCC 15439, which encode proteins
 CC AAY77190-Y77197.

XX

SQ Sequence 13842 BP; 1726 A; 5356 C; 4911 G; 1845 T; 4 other;

Query Match 69.2%; Score 16.6; DB 21; Length 13842;
 Best Local Similarity 82.6%; Pred. No. 2e+02;
 Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 ggggtcgacgtacgtcgagggggg 24

Db 6600 GGGCCGACGTCGCGAGGGGTG 6578

Search completed: August 10, 2002, 03:21:53
 Job time: 13684 sec

XX Pseudorabies virus; PRV; LTV; large latency transcript;
KW attenuated virus; vaccine; early protein 0; EP0; HSV-1 ICP0;
KW protecting animals; deletion mutants; swine; ds.
XX
OS Pseudorabies virus.
XX
XX
FH Key Location/Qualifiers
FT misc_feature 1..7013
FT FT /*tag= a
FT FT /note= "derived from PRV strain InPh"
FT FT 7014..8425
FT FT /*tag= b
FT FT /note= "derived from PRV strain Ka"
FT FT 622..6498
FT FT /*tag= c
FT FT /note= "encodes predicted amino acid sequence of ORF2"
FT FT 1..6
FT FT TATA_signal
FT FT /*tag= d
FT FT misc_feature 34
FT FT /*tag= e
FT FT /note= "RNA cap site"
FT FT 8382..8387
FT FT /*tag= f
XX US5352596-A.
XX
XX 04-OCT-1994.
XX
XX 11-SEP-1992; 92US-0945283.
XX
XX 11-SEP-1992; 92US-0945283.
XX
XX (USDA) US SEC OF AGRIC.
XX
XX Cheung AK, Wesley RD;
XX
XX WPI; 1994-316187/39.
XX P-PSDB; AAR60620.
XX
XX New pseudorabies virus mutants for use in vaccine - having a
XX deletion and/or insertion in the early protein 0 gene or large
XX latency transcript gene
XX
XX Disclosure; Column 15-30; 43pp; English.
XX
XX AAQ73500 shows the Pseudorabies virus (PRV) large latency transcript
XX (LTV). The basic sequence is derived from PRV strain InPh and PRV
XX strain Ka. The LTV overlaps and is transcribed in the opposite
XX orientation with respect to the EP0 (early polypeptide 0) and the
XX immediately early gene (IE180). EP0 is nonessential for replicatio,
XX LTV is the only gene expressed during PRV latency, and the IE180
XX gene is absolutely necessary for PRV replication. However there are
XX 2 copies of IE180 in the genome. It is expected that PRV lacking one
XX of the IE180 copies is viable. Deletions in the non-overlapping
XX regions of these 3 genes will generate single deletion routants,
XX while deletions in overlapping regions will generate double deletion
XX mutants. The invention is concerned with the construction of attenuated
XX viruses which have a reduced ability to reactivate from latency. This
XX can be achieved by functionally disabling the expression of the EP0
XX gene, or by disrupting the synthesis of the LTV, or both. (See also
XX AAQ73501 and AAR60620-24)
XX
XX Sequence 8438 BP; 1141 A; 2916 C; 3327 G; 1054 T; 0 other;
XX
XX
XX Query Match 69.2%; Score 16.6; DB 15; Length 8438;
XX Best Local Similarity 82.6%; Pred. NO. 2.1e+02;
XX Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX 1 ggggtcgcacgtactcgcaggagg 23
XX
XX 1315 GGGGGGACGCGTGTGACGGGG 1293

RESULT 15
AAZ87297/c
XX ID AAZ87297 standard; DNA; 13842 BP.
XX
XX AAZ87297;
XX
XX 05-JUN-2000 (first entry)
XX
XX S. venezuelae macrolide biosynthetic gene pikAI, SEQ ID NO:30.
XX
XX Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
XX neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
XX biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma,
XX chronic obstructive pulmonary disease; respiratory inflammation;
XX hypercholesterolaemia; crop protection agent; ds.
XX
XX Streptomyces venezuelae ATCC15439.
XX
XX Key Location/Qualifiers
XX CDS 1..13842
XX FT /*tag= a "pikAI"
XX FT /product= "pikAI"
XX FT /transl_except= (pos:4156..4158, aa:Ala)
XX FT /transl_except= (pos:13741..14743, aa:Ala)
XX
XX WO200000620-A2.
XX
XX 06-JAN-2000.
XX
XX 25-JUN-1999; 99WO-US14398.
XX
XX 26-JUN-1998; 98US-0105537.
XX
XX (MINU) UNIV MINNESOTA.
XX
XX Sherman DH, Liu H, Xue Y, Zhao L;
XX WPI; 2000-160679/14.
XX P-PSDB; AAY77192.
XX
XX Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.
XX synthesis of methymycin and pikromycin -
XX
XX Claim 15; Page 377-383; 438pp; English.
XX
XX The invention relates to an isolated and purified nucleic acid segment
XX comprising a desosamine biosynthetic gene cluster, a fragment or its
XX biologically active variant, where the nucleic acid sequence is not
XX derived from the eryC gene cluster of Saccharopolyspora erythraea or
XX Streptomyces antibioticus. The invention also relates to a macrolide
XX biosynthetic gene cluster, or fragments thereof. The macrolide
XX biosynthetic gene cluster encodes proteins which synthesize methymycin,
XX pikromycin, neomethymycin, narbomycin or a combination of these
XX compounds. Recombinant or augmented cells comprising the desosamine
XX and/or macrolide biosynthetic gene clusters are useful for the
XX production of biologically active macrolides. The macrolide biosynthetic
XX proteins are useful for synthesis of methymycin, pikromycin,
XX neomethymycin and narbomycin. The alternative termination of polyketide
XX synthesis may be useful to prepare novel antibiotics and
XX polyhydroxyalkanoate (PHA) monomers. The compounds produced by the
XX recombinant host cells are useful as biopolymers, e.g., in packaging or
XX biomedical applications, to engineer PHA monomer synthases or to prepare
XX biologically active agents, such as chemotherapeutics,
XX immunosuppressants, agents to treat asthma, chronic obstructive pulmonary
XX disease as well as other diseases involving respiratory inflammation,
XX cholesterol-lowering agents or macrolide-based antibiotics which are
XX active against a variety of organisms, e.g., bacteria, including
XX multi-drug resistant pneumococci and other respiratory pathogens, as well
XX as viral parasitic pathogens, or as crop protection agents (e.g.,
XX fungicides or insecticides) via expression of polyketides in plants.
XX
XX Sequences AAZ87295-287302 represent macrolide biosynthetic genes from

Query Match 69.2%; Score 16.6; DB 21; Length 4689;
Best Local Similarity 82.6%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 gggtcgacgtacgtcgagggggg 24
||| ||||| ||| ||||| |||
DB 3324 GGGCCGACGTCGTCGGAGGGGTG 3302

RESULT 12
ABL33599
ID ABL33599 standard; DNA; 5487 BP.
XX
AC ABL33599;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1572.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PT 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 1572; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 5487 BP; 1608 A; 133 C; 1240 G; 2506 T; 0 other;

Query Match 69.2%; Score 16.6; DB 24; Length 5487;
Best Local Similarity 82.6%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 gggtcgacgtacgtcgagggggg 24
||| ||||| ||| ||||| |||
DB 616 gggtcgacgtacgtcgagggggg 638

RESULT 13
ABL34148
ID ABL34148 standard; DNA; 6242 BP.
XX
AC ABL34148;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 2121.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PT 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 2121; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 6242 BP; 1699 A; 148 C; 1349 G; 3046 T; 0 other;

Query Match 69.2%; Score 16.6; DB 24; Length 6242;
Best Local Similarity 82.6%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 gggtcgacgtacgtcgagggggg 24
||| ||||| ||| ||||| |||
DB 4883 gcgtcgcgtaagtcgagggcgg 4905

RESULT 14
AAQ73500/c
ID AAQ73500 standard; DNA; 8438 BP.
XX
AC AAQ73500;
XX
DT 15-MAY-1995 (first entry)
XX
DE DNA encoding Pseudorabies virus large latency transcript.

```
RESULT 10
AAS90594/C
ID AAS90594 standard; cDNA; 2027 BP.
XX
AC AAS90594;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #26398.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ABG26407.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 1; SEQ ID No 26398; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome.
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 2027 BP; 422 A; 574 C; 581 G; 450 T; 0 other;

Query Match 69.2%; Score 16.6; DB 23; Length 2027;
Best Local Similarity 82.5%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gggggtcagctacgtcaggggg 23
      ||||| ||| ||| ||| |||
DB 821 GGGGTCGGCGTGGTGGAGCGG 799

RESULT 11
AAZ87299/C
ID AAZ87299 standard; DNA; 4689 BP.
XX
XX AAZ87299;
XX
XX 05-JUN-2000 (first entry)
XX
XX S. venezuelae macrolide biosynthetic gene pikAIII, SEQ ID NO:34.
XX
XX Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
XX neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
XX biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma,
XX chronic obstructive pulmonary disease; respiratory inflammation;
XX hypercholesterolaemia; crop protection agent; ds.
XX
XX Streptomyces venezuelae ATCC15439.
XX
XX Key Location/Qualifiers
XX CDS 1..4689
XX FT /*tag= a
XX FT /product= "PikAIII"
XX
XX WO200000620-A2.
XX
XX 06-JAN-2000.
XX
XX 25-JUN-1999; 99WO-US14398.
XX
XX 26-JUN-1998; 98US-0105537.
XX
XX (MINU ) UNIV MINNESOTA.
XX
XX Sherman DH, Liu H, Xue Y, Zhao L;
XX
XX WPI; 2000-160679/14.
XX
XX P-PSDB; AAY77194.
XX
XX Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.
XX synthesis of methymycin and pikromycin -
XX
XX Claim 15; Page 415-417; 438pp; English.
XX
XX The invention relates to an isolated and purified nucleic acid segment
XX comprising a desosamine biosynthetic gene cluster, a fragment or its
XX biologically active variant, where the nucleic acid sequence is not
XX derived from the eryC gene cluster of Saccharopolyspora erythraea or
XX Streptomyces antibioticus. The invention also relates to a macrolide
XX biosynthetic gene cluster, or fragments thereof. The macrolide
XX biosynthetic gene cluster encodes proteins which synthesise methymycin,
XX pikromycin, neomethymycin, narbomycin or a combination of these
XX compounds. Recombinant or augmented cells comprising the desosamine
XX and/or macrolide biosynthetic gene clusters are useful for the
XX production of biologically active macrolides. The macrolide biosynthetic
XX proteins are useful for synthesis of methymycin, pikromycin,
XX neomethymycin and narbomycin. The alternative termination of polyketide
XX synthesis may be useful to prepare novel antibiotics and
XX polyhydroxyalkanoate (PHA) monomers. The compounds produced by the
XX recombinant host cells are useful as biopolymers, e.g., in packaging or
XX biomedical applications, to engineer PHA monomer synthetases, or to prepare
XX biologically active agents, such as chemotherapeutics,
XX immunosuppressants, agents to treat asthma, chronic obstructive pulmonary
XX disease as well as other diseases involving respiratory inflammation,
XX cholesterol-lowering agents or macrolide-based antibiotics which are
XX active against a variety of organisms, e.g., bacteria, including
XX multi-drug resistant pneumococci and other respiratory pathogens, as well
XX as viral parasitic pathogens, or as crop protection agents (e.g.,
XX fungicides or insecticides) via expression of polyketides in plants.
XX Sequences AAZ87295-287302 represent macrolide biosynthetic genes from
XX Streptomyces venezuelae ATCC 15439, which encode proteins
XX AAY77190-Y77197.
XX
XX Sequence 4689 BP; 648 A; 1882 C; 1572 G; 587 T; 0 other;
```

XX WPI; 1994-316187/39.
DR P-PSDB; AAR60621.
XX
PT New pseudorabies virus mutants for use in vaccine - having a
PT deletion and/or insertion in the early protein O gene or large
PT latency transcript gene
XX
XX Disclosure; Column 39-44; 43pp; English.
PS
PS AAQ73501 shows the DNA sequence of the early polypeptide 0 (EP0) gene.
CC EP0 is nonessential for replication, LIT (large latency transcript) is
CC the only gene expressed during PRV latency, and the IE180 gene is
CC absolutely necessary for PRV replication. However there are 2 copies of
CC IE180 in the genome. It is expected that PRV lacking one of the IE180
CC copies is viable. Deletions in the non-overlapping regions of these 3
CC genes will generate single deletion routants, while deletions in
CC overlapping regions will generate double deletion mutants. The invention
CC is concerned with the construction of attenuated viruses which have a
CC reduced ability to reactivate from latency. This can be achieved by
CC functionally disabling the expression of the EP0 gene, or by disrupting
CC the synthesis of the LIT, or both. (See also AAQ73500 and AAR60620-24)
XX
SQ Sequence 1682 BP; 256 A; 619 C; 550 G; 257 T; 0 other;

Query Match 69.2%; Score 16.6; DB 15; Length 1682;
Best Local Similarity 82.6%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcgcgtacgtcgcgagggg 23
||||| ||||| | ||||| |||||
Db 531 gggggcgcgaggtcgcgagggg 553

RESULT 8
AAQ10213/C
ID AAQ10213 standard; DNA; 1831 BP.
XX
AC AAQ10213;
XX
DT 17-DEC-2001 (updated)
DT 27-MAR-1991 (first entry)
XX
DE BamHI G-P-J fragment carrying sequences characteristic of latent
DE pseudorabies virus.
XX
KW PRV; ss.
XX
OS Pseudorabies virus.
XX
PN USN7537855-N.
XX
PD 18-DEC-1990.
XX
PF 13-JUN-1990; 90US-0238940.
XX
PR 13-JUN-1990; 90US-0537855.
XX
PA (USDA) US AGRIC RES SERV.
XX
PI Cheung AK;
XX
DR WPI; 1991-021957/03.
XX
PT Pseudo-rabies virus nucleotide sequences - used for producing
PT nucleic acid probes, antigens and antibodies for distinguishing
PT latent from productive infection
XX
PS Disclosure; Page 22; 27pp; English.
XX
XX The fragment carries sequences characteristic of the latent
CC pseudorabies viral genome, and may be used as a probe in diagnosis

CC of infection.
CC (Note: Revised entry submitted to correct the patent number format of
CC US Government-owned NTIS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpi/updates/ntis_us.html.)
XX
SQ Sequence 1831 BP; 305 A; 643 C; 639 G; 244 T; 0 other;

Query Match 69.2%; Score 16.6; DB 12; Length 1831;
Best Local Similarity 82.6%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcgcgtacgtcgcgagggg 23
||||| ||||| | ||||| |||||
Db 1317 GGGGCGGACGGATGTCGACGGGG 1295

RESULT 9
AAQ10211/C
ID AAQ10211 standard; DNA; 1831 BP.
XX
AC AAQ10211;
XX
DT 17-DEC-2001 (updated)
DT 27-MAR-1991 (first entry)
XX
DE BamHI G-P-J fragment carrying sequences characteristic of productive
DE pseudorabies virus.
XX
KW PRV; ss.
XX
OS Pseudorabies virus.
XX
PN USN7537855-N.
XX
PD 18-DEC-1990.
XX
PF 13-JUN-1990; 90US-0238940.
XX
PR 13-JUN-1990; 90US-0537855.
XX
PA (USDA) US AGRIC RES SERV.
XX
PI Cheung AK;
XX
DR WPI; 1991-021957/03.
XX
PT Pseudo-rabies virus nucleotide sequences - used for producing
PT nucleic acid probes, antigens and antibodies for distinguishing
PT latent from productive infection
XX
PS Disclosure; Page 20; 27pp; English.
XX
XX The fragment carries sequences characteristic of the productive
CC pseudorabies viral genome, and may be used as a probe in diagnosis
CC of infection.
CC (Note: Revised entry submitted to correct the patent number format of
CC US Government-owned NTIS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpi/updates/ntis_us.html.)
XX
SQ Sequence 1831 BP; 305 A; 643 C; 639 G; 244 T; 0 other;

Query Match 69.2%; Score 16.6; DB 12; Length 1831;
Best Local Similarity 82.6%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcgcgtacgtcgcgagggg 23
||||| ||||| | ||||| |||||
Db 1317 GGGGCGGACGGATGTCGACGGGG 1295

CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SQ Sequence 24 BP; 3 A; 4 C; 14 G; 3 T; 0 other;

Query Match 100.0%; Score 24; DB 22; Length 24;

Best Local Similarity 100.0%; Pred. No. 0.17; Mismatches 0; Gaps 0;

Matches 24; Conservative 0; Indels 0; Gaps 0;

Qy 1 ggggtcgacgtacgtcgagggggg 24

Db 1 ggggtcgacgtacgtcgagggggg 24

RESULT 4

AXX22281

ID AAX22281 standard; DNA; 1327 BP.

XX

AC AAX22281;

XX

DT 18-MAY-1999 (first entry)

XX

DE Human secreted protein gene 63 clone HJAAAT30.

XX

KW Human: secreted protein; gene therapy; protein therapy; cancer; weight;
KW tumour; chromosome mapping; forensic; haematological disease; allergy;
KW inflammation; cell proliferation; viral infection; wound healing;
KW modulation; appetite; behaviour; food additive; preservative; ss.

XX

OS Homo sapiens.

XX

PN WO9903990-A1.

XX

PD 28-JAN-1999.

XX

PF 15-JUL-1998; 98WO-US14613.

XX

PR 18-AUG-1997; 97US-0056361.

PR

PR 16-JUL-1997; 97US-0052661.

PR

PR 16-JUL-1997; 97US-0052870.

PR

PR 16-JUL-1997; 97US-0052871.

PR

PR 16-JUL-1997; 97US-0052872.

PR

PR 16-JUL-1997; 97US-0052873.

PR

PR 16-JUL-1997; 97US-0052874.

PR

PR 16-JUL-1997; 97US-0052875.

PR

PR 22-JUL-1997; 97US-0053440.

PR

PR 22-JUL-1997; 97US-0053441.

PR

PR 22-JUL-1997; 97US-0053442.

PR

PR 18-AUG-1997; 97US-0055683.

PR

PR 18-AUG-1997; 97US-0055724.

PR

PR 18-AUG-1997; 97US-0055725.

PR

PR 18-AUG-1997; 97US-0055726.

PR

PR 18-AUG-1997; 97US-0055946.

PR

PR 18-AUG-1997; 97US-0055952.

PR

PR 18-AUG-1997; 97US-0055985.

PR

PR 18-AUG-1997; 97US-0055989.

PR

PR 18-AUG-1997; 97US-0056359.

PR

XX (HUMA-) HUMAN GENOME SCI INC.

XX

XX Duan R, Feng P, Ferrie AM, Florence KA, Fouad J;

XX Greene JM, Hu J, Ni J, Rosen CA, Ruben SM, Young PB;

XX Yu G;

XX WPI; 1999-132234/11.

XX P-PSDB; AAY01453.

XX

XX

PT

PT

PT

XX

XX

XX

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Query Match 73.3%; Score 17.6; DB 20; Length 1327;

Best Local Similarity 83.3%; Pred. No. 83;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ggggtcgacgtacgtcgagggggg 24

Db 51 ggggtcgacgtacgtcgagggggg 74

RESULT 5

ABL34157

ID ABL34157 standard; DNA; 16766 BP.

XX

AC ABL34157;

XX

DT 26-MAR-2002 (first entry)

XX

DE Human immune system associated gene SEQ ID NO: 2130.

XX

XX Human: immune system disease; cytosine methylation; antiasthmatic;

XX antiarteriosclerotic; antianemic; cytostatic; neurotropic;

XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;

XX antirheumatic; antiarthritic; antidiabetic; antipsoriatic;

XX antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;

XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;

XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;

XX gene; ds.

XX

OS Homo sapiens.

XX

XX WO200200928-A2.

PN

XX 03-JAN-2002.

PD

XX 02-JUL-2001; 2001WO-EP07537.

PF

XX 30-JUN-2000; 2000DE-1032529.

PR

New nucleic acids encoding secreted human proteins - potentially
useful for treating and diagnosing diseases and identifying specific
binding agents

Claim 4; Page 205; 251pp; English.

The invention relates to nucleic acid sequences (AAX22211 to AAX22282)
encoding human secreted proteins (AAY01383 to AAY01454). The secreted
protein gene sequences are deposited with the ATCC under deposit number
ATCC 209138, 209139 or 209141. Host cells containing vectors comprising
the nucleic acid sequences are used for the recombinant expression of
the secreted proteins. The polynucleotide and amino acid sequences are
useful for preventing, treating or ameliorating medical conditions e.g.
by protein or gene therapy. Pathological conditions can be also diagnosed
by determining the amount of the new polypeptides in a sample or by the
presence of mutations in the new polynucleotides. The nucleic acid
sequences, or its fragments, are useful for chromosome identification
and mapping; as antisense and triplex-forming therapeutics; in gene
therapy; for (forensic) identification of individuals; as molecular
weight markers; to identify related sequences or specific mRNA; in
preparation of oligomers and to raise anti-DNA antibodies. Antibodies are
useful as immunoassay reagents (including for in vivo imaging) and
therapeutically to inhibit or activate particular polypeptides. A very
wide range of disorders may be treated with the polynucleotide and
polypeptide sequences, e.g. autoimmune or haematological diseases,
allergy, inflammation, cancer or other forms of cell proliferation, viral
or other infections. The sequences may also be useful in wound healing,
to modulate differentiation of embryonic stem cells, to modulate weight,
appetite, behaviour etc. and as food additive or preservative. The
present sequence represents a gene encoding a human secreted protein
(see descriptor line for gene number and clone identification).

Sequence 1327 BP; 386 A; 258 C; 298 G; 359 T; 26 other;


```

CDS
join(13372..13429,13884..13952,15928..16063,17127..17291,
17806..17911,17998..18061,18385..18492,18716..18865,
19017..19271,19540..19745,19949..20176,22473..22498,
22949..23016,24394..24532,24781..24889,25112..25222)
/gene="P0583608.4"
/note="hypothetical protein
similar to Arabidopsis thaliana protein F28A23.100"
/codon_start=-1
/protein_id="BAB64770.1"
/db_xref="GI:15528704"
/translation="MVEHNGWGRSSVYRCMAGCGGRIAMQFOLRLVRRONQCP
QESAFYRRAFIKRYTHADELGLKLEADVRKSDNTIKRSFRVAGEMRRARPCSLSE
SPVLLRRRAAMOGGGDADVEASGARFEDDESQLYHAHSTGFYHDPAVAGWYSGRD
GOYVYLDNGNYLWPTPDAGNEPXTNHPNCDENODPLEPSSGLEODIPDNQETORPS
EMLLKNSMLLPDEFSEYGDSPSETAENKPCNFASENAPSSPCASLHOTDELIQENS
TAVHESLSEEEKWLAQGOVDEIDDPLEPIDLMDMWKSEVSKGQPMARLVR
LYKSSKPHPSLPARGGLRTAPVHEVHLDIVAVSSGKITRLANPSRKTIASISYDS
SNPTKDMGPNITYANSNDNDLKQSAQCOSEVDFVSMKGVSAASAKKHTINAYRDA
AERRLIHRGIGIPGQKQSSSTNFEHEETSECDMDLGSVDMNFRSSGLNKKMM
ENMGKERKLLIVSKNLNCKGLVTEEREKIVNIPDNGYGLSDRIILGFGYGYE
YGFELPDTNMTMNSGFPSPDNNSDIRKICFRIRVRSYSILIRIRDNPLYFPYSYVY
SPSYPOOVVDEIAHQTETVGAHHTMOIKTICHTSTSSNCP"
join(29003..31883,32004..32468,33804..33994)
/gene="P0583608.5"
join(29003..31883,32004..32468,33804..33994)
/gene="P0583608.5"
/note="contains EST C74881(BE6016)"
/codon_start=-1
/product="putative receptor protein kinase"
/protein_id="BAB64771.1"
/db_xref="GI:15528705"
/translation="MPPGLSRRRRPPAALLCCVAVVAACWVGALADAOAGALLA
MKTELRGDTALPDMPADAPSCRTMVGRCNANGRTETELSOVDLGGVPMLSAM
GTLIELVLGAGNLGSPITPAOLGIDPALHLDLSNATLGLSPISALCSGKESLIVY
NSMHLGATIPDAGNLTALRELIIPNODGAPATIGOMASIEVIRGGCKNLQAL
PEPTGCSKILTMGLAETISGCPPLATDQKUNLTAITLALISGPIPELRCISL
ENIYIENALSGSTPAQDGLANLKNLMLNOMNLVNIPELACGLAVVDLSMGL
TGHPASLNLSSLDLQLSVNKGSPITAEISRCNLDLEDDNOIGALIPAEIGK
LYALMLLYMANOLTGITPEIGCAGLSLSDSONALGPIRISFLRPLSKLLI
DNTLGEIPPEIGNCTSLVRPGSLNHLGNDIPPEGKISFELSLSTRSGALIP
ETAGCNLTFVLDHGNALGVLPRGLFQCPSPLOYDLSYNAGALIPANIGMSIT
KVLGNGRLSGQIPPEIGCSRLQDLDSNSTGATITMSIGTIPCLLALNLSGCL
SGAIPKGFAGLAVLGVNDSHNOITGDLPPLSLQNLVNLISYNFTGRAPETAFA
RLPASVEGNPGICLSRCPGDASDRERRARRARVAVLSSVALLAAAEVLEFR
RROPLFRGSTPADGDKDADMLPMDVTLVOKLEISYGDVARSILTPANVIGQWGS
AVYRASIPSTGVAIAVKKRRSDSEAVDAFACVEGLVPRVHRNIVRLGEMANRRT
LLEFDYLPNGTIGLLHGGGAALGAIVEMEVRLSTAVGAEGLAYLHDSVPAILHR
DYKSDNIIIGERYEACIADPGLARVADDCANSPPPASGYCIAPAYECMKITTKS
DYISGCVLLEITIGRRPIEAFEGEGQIVYVWYREHLHRKRDPAEYIDSLQGRSDTO
VOEMLQALGIALCASTRPDRPTMRDVALLGLRHDSAEQKAGSGSAIKMADPR
QSGSPTKPMQAQAAHSHTSLSLAYSTGSPGEPPASPPAPKDPNPNSAAGKRAAGCG
PAPWSWSSPTKTTAARTEQSSVYSSSLVRLF"
34110..36208
/gene="P0583608.6"
/note="pseudogene"
/pseudo
34110..36208
/gene="P0583608.6"
complement(join(41292..41669,42043..42189,42758..42811,
44914..45078,45669..45968,46090..46176,46326..46502))
/gene="P0583608.7"
complement(join(41292..41669,42043..42189,42758..42811,
44914..45078,45669..45968,46090..46176,46326..46502))
/note="hypothetical protein
similar to Arabidopsis thaliana chromosome 5, AAP04872.1"
/codon_start=-1
/protein_id="BAB64772.1"
/db_xref="GI:15528706"
/translation="MASSGCSNRLRQTIIVALLLSASSCLASAKHRGNGTTTVPF
HGKDELRAVRKIMAOSSPDGDIVDVCVAHLOPAFEHPIKLGOKPEAREPERPKVGA
AAAEERAVYPOAMTQGESCEPCTVYVRRTRRDVLRSSSVAFRGKQKOPRAAGVYRD
STSDGHEVSQSTNSPVTGELAVGKIALHHAIVGIVTGDQFYGAASLNVASARVATAE
```

```

FSLQIWTWISGSRNDLNTIAGMOVSPELYGDNDRFFTYWTTDAYQATGCVNLHCS
GFVQINRIALIGAALISTSVYNGRCIDILLIMKDRRGRHMMIQLSGDLGVWPSL
FTLLHGAMNVQVSGEVNTRPNSGSHTPMOMSGHPRRGCFNAAAFERMLQVYDMDN
LIPAAALRIYADHPSCYDIOGCYNNRMNGVFFYGGGRVRCPC"
join(53067..53187,53300..53397,53692..53733,54064..54278,
54455..54502,55438..55574,56735..56928,57099..57208,
57321..57439,57764..57842)
/gene="P0583608.8"
join(53067..53187,53300..53397,53692..53733,54064..54278,
54455..54502,55438..55574,56735..56928,57099..57208,
57321..57439,57764..57842)
/note="hypothetical protein
similar to Arabidopsis thaliana chromosome 3, AAP26101.1"
/codon_start=-1
/protein_id="BAB64773.1"
/db_xref="GI:15528707"
/translation="MAWRGLARDAAAYLRRCAPAHVFSASRAPGVANPARGFR
RMGLVPEPPELSLSSGMFVDSIRSCQVAYAEVAPQSEGLGPSTKHQISKLW
TIRKYLQVGLALIALOMQDGLIFIVLLILYSSRSPYSIYLFQEVHGEHQ
NRAFEREAVLFRKVDTKDKLPSTIGTSADREVLHVIGILGMMIYRASGKEONL
FVWGKRTIVAVDFSSCKRALRMATNLFRRGDDVLITHNSVYHNEGAYOIMPOS
GSFLPLAEFSDBPHAKTYAVSPDKETLEILNOMSKQREVLAKILYDPAKKLIEA
VDLVPNLCLVGNRGISTLRALMGSSVYVNNATCPVTYKENT"
join(59847..60174,60209..60263,60286..60574,61172..61324,
62306..63169,63296..63409,64875..65140,65697..65993,
66434..66804,67702..67758,68058..68086,68430..70194,
70984..71093,71170..71237,71444..71558)
/gene="P0583608.9"
join(59847..60174,60209..60263,60286..60574,61172..61324,
62306..63169,63296..63409,64875..65140,65697..65993,
70984..71093,71170..71237,71444..71558)

```

Search completed: August 10, 2002, 02:58:57
Job time: 15743 sec

```

Query Match      76.3%: Score 20.6; DB 8; Length 135285;
Best Local Similarity 85.2%: Mismatch No. 1, 7e+03;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 999gtcgagctcgagctcgagggggg 27
Db 68926 GGGGTCGACGACGACGAGAGGCG 68900
```

Sat Aug 10 09:08:34 2002

us-09-672-126-36.rge

Page 15

THIS PAGE BLANK (USPTO)

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	27	100.0	27	22	AAF98766	Human IFN- α .im
2	27	100.0	27	22	AAF99871	Immunostimulatory
3	21.4	79.3	1734	16	AA294584	Maize cyclin D zmc
C 4	20.6	76.3	1438	16	AAT05310	partial human fibr
C 5	20.6	76.3	1438	21	AAK57835	CDNA encoding the
C 6	20.6	76.3	7379	19	AAV13176	Complete DNA sequ
7	20.6	76.3	23914	23	ABU20258	Drosophila melanog
8	20.2	74.8	7317	22	AAK4534	Chemically pretrea
C 9	19.8	73.3	2032	19	AAV26974	Consensus I3L prom

FK 2/-SEP-1993, 3303-0130147.

ALIGNMENTS

RESULT	1	
AAF98766		
ID	AAF98766 standard; DNA; 27 BP.	
XX		
XX	AAF98766;	
XX		
XX	11-JUN-2001 (first entry)	
XX		
XX	Human IFN-alpha Immunostimulatory nucleic acid SEI	
DE		
DE	Immunostimulatory nucleic acid; ISNA; human; inte	
KW	viral infection; phosphorothioate backbone; palin	
KW		
XX	Synthetic.	
XX		
Key	Location/Qualifiers	
modified_base	1..2	
FT	./tag= a	
FT	/mod_base= "OTHER"	
FT	/note= "phosphorothioate linkage"	
modified_base	22..26	
FT	./tag= b	
FT	/mod_base= "OTHER"	
FT	/note= "phosphorothioate linkage"	
XX		
XX		
PN	WO200122990-A2.	
XX		
XX		
PD	05-APR-2001.	
XX		
PF	27-SEP-2000; 2000WO-US26527.	
XX		
PR	27-SEP-1999; 99US-0156147.	

XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX
 PI Hartmann G, Bratzler RL, Krieg A;
 XX WPI; 2001-290487/30.
 DR
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 XX
 PS Claim 201; Page 103; 168pp; English.
 XX
 CC The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX
 SQ Sequence 27 BP; 3 A; 5 C; 16 G; 3 T; 0 other;
 Query Match 100.0%; Score 27; DB 22; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.65;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ggggtcgacgtcgacgtcgagggggg 27
 Db 1 ggggtcgacgtcgacgtcgagggggg 27
 RESULT 2
 AAF99871
 ID AAF99871 standard; DNA: 27 BP.
 XX
 AC AAF99871;
 XX
 DT 12-JUN-2001 (first entry)
 XX
 DE Immunostimulatory nucleic acid #987.
 XX
 KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX
 OS Synthetic.
 XX
 PN WO200122972-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 25-SEP-2000; 2000WO-US26383.
 XX
 PR 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX
 PA (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 XX
 PI Krieg AM, Schetter C, Vollmer J;
 XX WPI; 2001-273485/28.
 XX
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 PT
 XX Claim 101; Page 59; 338pp; English.

XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 XX
 SQ Sequence 27 BP; 3 A; 5 C; 16 G; 3 T; 0 other;
 Query Match 100.0%; Score 27; DB 22; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.65;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ggggtcgacgtcgacgtcgagggggg 27
 Db 1 ggggtcgacgtcgacgtcgagggggg 27
 RESULT 3
 AAZ94584
 ID AAZ94584 standard; DNA: 1734 BP.
 XX
 AC AAZ94584;
 XX
 DT 18-JUL-2000 (first entry)
 XX
 DE Maize cyclin D ZmCycD gene.
 XX
 KW Maize; cyclin D; ZmCycD gene; CycD; cell division; cell cycle;
 KW transgenic plant; ss.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 FF CDS 213..1262
 FT /*tag= a
 FT
 XX WO200017364-A2.
 PN
 XX 30-MAR-2000.
 PD
 XX 21-SEP-1999; 99WO-US21946.
 PF
 XX 23-SEP-1998; 98US-0101551.
 PR
 XX (PION-) PIONEER HI-BRED INT INC.
 PA
 XX Lowe KS, Tao Y, Gordon-Kamm WJ, Gregory CA, McElver JA;
 PI Hoerster GJ;
 PI
 DR WPI; 2000-283589/24.
 DR P-PSDB; AA79324.
 XX
 PT Novel polynucleotides encoding maize cyclin D isoforms 1, 2 and 3,
 PT related proteins and antisense RNA useful for control of cell cycle
 PT regulation -
 XX
 PS Claim 1; Page 126-128; 134pp; English.
 CC
 CC The present sequence is that of an isoform of the maize ZmCycD
 CC gene that encodes cyclin D (CycD, see AA79324), a protein necessary
 CC for progression from G1 into S phase. The encoded protein binds to
 CC CDK4, and the active CycD-CDK4 hyperphosphorylates retinoblastoma

CC associated protein, releasing the E2F transcription factor which
 CC activates DNA synthesis. The invention provides maize CycD
 CC polynucleotides (see AAZ94581-84) and polypeptides (see AAY79321-24)
 CC that are involved in cell cycle regulation. Also provided are
 CC recombinant expression cassettes (including 2mCycD in sense or
 CC antisense orientation), host cells, transgenic plants (especially
 CC corn, sorghum, sunflower, safflower, wheat, rice, alfalfa or
 CC oilseed Brassica) and antibody compositions. A claimed method of
 CC modulating the level of CycD protein in a cell comprises
 CC transforming the cell with a recombinant expression cassette
 CC comprising a CycD polynucleotide linked to a promoter, and
 CC growing the cell for a time sufficient to induce expression of the
 CC polynucleotide sufficient to modulate (increase or decrease) the
 CC CycD protein in the cell. The CycD protein is present in an amount
 CC sufficient to alter cell division, increase the number of cells
 CC dividing, improve transformation frequencies, alter cell growth,
 CC increase the growth rate, increase crop yield, alter plant
 CC height or size, enhance or inhibit organ (seed, root, shoot, ear,
 CC tassel, stalk, pollen, stamen) growth, produce organ ablation,
 CC produce parthenocarpic fruits, produce male sterile plants,
 CC enhance embryogenic response, increase callus induction, provide
 CC positive selection, increase plant regeneration, alter the time
 CC that cells are arrested in G1 or G0 phase or in a particular cell
 CC cycle, improve response to environmental stress including
 CC dehydration, heat or cold, increase the number of pods per plant,
 CC increase the number of seeds per pod or ear, alter the lag time in
 CC seed development, provide hormone-independent cell growth, or
 CC increase the growth rate of cells in bioreactors. The level of
 CC CycD protein in the cells is transiently modulated by introducing
 CC CycD RNA or CycD polypeptides. CycD polynucleotides can be used
 CC to identify CycD interacting proteins. All claimed.
 XX
 SQ Sequence 1734 BP; 321 A; 519 C; 534 G; 360 T; 0 other;

Query Match 79.3%; Score 21.4; DB 21; Length 1734;
 Best Local Similarity 95.7%; Pred. No. 42;
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 gtcagctcgacgtcgagggggg 26
 |||||
 Db 1070 gtcagctcgacgtcgagggggg 1092

RESULT 4
 AAT05310/C
 ID AAT05310 standard; DNA; 1438 BP.
 XX
 AC AAT05310;
 XX
 DT 12-JUN-1996 (first entry)
 XX
 DE Partial human fibrinogen gamma-chain DNA.
 XX
 KW Human fibrinogen; gamma-chain; synthetic 3'-end fragment;
 KW Bluescript II KS+; plasmid; mp19gamma2; expression vector; PREP9;
 KW variant fibrin chains; unable to self polymerise; fibrinogen;
 KW surgical sealants; thrombin activation; pure starting material;
 KW fibrin-derived factors; regulation; angiogenesis;
 KW platelet aggregation; ds.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX CDS 3..1367
 XX FT /*tag= a
 XX FT /note= "START codon absent"
 XX
 XX W09529686-A1.
 XX
 XX 09-NOV-1995.
 XX
 XX 02-MAY-1995; 95WO-US05527.
 XX
 XX

XX 02-MAY-1994; 94US-0236979.
 PR (SQUI) SQUIBB & SONS INC E R.
 PA Cederholm-Williams SA;
 PI WPI; 1995-392917/50.
 DR P-PSDB; AAR84551.
 DR
 XX Variant chains of fibrin unable to self polymerise - are able to
 PT react with fibrinogen, partic. usefull in surgical sealants that do
 PT not require activation of thrombin
 PS Disclosure; Fig 7; 102pp; English.
 XX
 CC Using the primers AAT05292/93 a 310 bp fragment from a human
 CC fibrinogen gamma-chain cDNA clone was amplified, and digested to
 CC allow the N-terminal and C-terminal portions of the gamma-chain
 CC (a partial nucleotide and amino acid sequence of which is given
 CC in AAT05310 and AAR84551, respectively) to be purified. They were
 CC then ligated along with the synthetic 3'-end fragment AAT05309,
 CC and cloned into a mp19 vector to give mp19gamma2, which encodes
 CC a complete gamma-chain. mp19gamma2 was then subcloned into the
 CC expression vector PREP9, which was used in the prodn. of variant
 CC fibrin chains unable to self polymerise. These chains are able
 CC to react with fibrinogen, partic. usefull in surgical sealants
 CC that do not require thrombin activation, and are pure starting
 CC materials for fibrin-derived factors that regulate angiogenesis,
 CC platelet aggregation, etc..
 XX
 SQ Sequence 1438 BP; 454 A; 293 C; 322 G; 369 T; 0 other;

Query Match 76.3%; Score 20.6; DB 16; Length 1438;
 Best Local Similarity 85.2%; Pred. No. 81;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ggggtcgacgtcgacgtcgagggggg 27
 |||||
 Db 36 GGGGCCGCGTCTGACCTCGAGGGGGG 10

RESULT 5
 AAA57835/C
 ID AAA57835 standard; cDNA; 1438 BP.
 XX
 AC AAA57835;
 XX
 DT 20-OCT-2000 (first entry)
 XX
 DE cDNA encoding the gamma chain of human fibrinogen.
 XX
 KW Fibrin sealant; fibrin; surgery; bleeding; adhesion; surgical adhesive;
 KW angiogenesis; platelet aggregation; gamma-fibrinogen; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX CDS 3..1367
 XX FT /*tag= a
 XX FT /product= "gamma-fibrinogen"
 XX
 XX US6083902-A.
 XX
 XX 04-JUL-2000.
 PD
 XX 03-MAY-1995; 95US-0434099.
 XX
 XX 02-MAY-1994; 94US-0236979.
 XX
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX

PI Cederholm-williams SA;

XX WPI; 2000-464370/40.
 DR P-PSDB; AAY94009.

XX New fibrin sealant containing human fibrin homolog unable to
 PT self-polymerize but forms non-covalent bond with fibrinogen, useful in
 PT surgery for controlling bleeding or adhering tissues to each other -
 XX Example; Fig 7; 39pp; English.

XX The present sequence encodes the gamma chain of human fibrinogen.
 CC The sequence was used to produce a fibrin sealant. The specification
 CC describes a fibrin sealant which contains a human fibrin-homologue. The
 CC sealant comprises a recombinant variant fibrin chain differing from the
 CC naturally occurring gamma-chain by one or more mutations or deletions in
 CC a C-terminal region following a coiled-coil forming region. When
 CC incorporated into a fibrin-homologue, the homologue cannot self-
 CC polymerize but forms non-covalent bonds or polymerize with fibrinogen.
 CC The fibrin sealants are used in surgery to control bleeding or to adhere
 CC two tissues to each other. The recombinant fibrin chains are used may be
 CC used in the preparation of safe and convenient surgical adhesives and
 CC sealants, and as sources of substantially pure starting material for
 CC the production of fibrin-derived factors that regulate angiogenesis, or
 CC platelet aggregation. Fibrin and fibrin homologues may be used as
 CC components of fibrin monomer-based surgical sealants.

XX Sequence 1438 BP; 454 A; 293 C; 322 G; 369 T; 0 other;

Query Match 76.3%; Score 20.6; DB 21; Length 1438;
 Best Local Similarity 85.2%; Pred. NO. 81;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcagctcgacgtcgagggggg 27
 Db 36 GGGGCGCGGTGACCTCGAGGGGGG 10

RESULT

AAV13176/6
 ID AAV13176 standard; DNA; 7379 BP.

XX AAV13176;
 AC AAV13176;

XX 16-JUL-1998 (first entry)
 DT

XX Complete DNA sequence of plasmid pLF092.
 DE

XX Plasmid pLF092; canine adenovirus; vaccine; ss.
 KW
 XX - Synthetic.

XX WO9800166-A1.
 PN

XX 08-JAN-1998.
 PD

XX 30-JUN-1997; 97WO-US11486.
 PF

XX 03-JUL-1996; 96US-0675566.
 PR

XX 03-JUL-1996; 96US-0675556.
 PR

XX (INMR) RHONE MERIEUX INC.
 PA

XX Fischer L;
 PI

XX WPI; 1998-086736/08.
 DR

XX Canine adenovirus synthetically modified to contain exogenous DNA
 PT where non-essential region of virus has been deleted, useful in
 PT immunogenic, immunological or vaccine composition(s)
 PT Example 12; Fig 25; 226pp; English.

XX

PS

XX

CC The present plasmid relates to an invention where a canine
 CC adenovirus (cAd) is synthetically modified to contain exogenous DNA,
 CC where a non-essential region of the cAd has been deleted.

CC An immunogenic, immunological or vaccine composition comprising the
 CC the above cAd can be used to induce an immunological response in a
 CC host vertebrate, preferably a canine or human, to which it is
 CC administered, or transfer genetic information to an animal or
 CC human. The exogenous DNA preferably encodes an expression product
 CC comprising an epitope of interest, biological response modulator,
 CC growth factor, recognition sequence, therapeutic gene or fusion
 CC protein, e.g. a Morbillivirus antigen, rabies glycoprotein, tumour
 CC necrosis factor or melanoma associated antigen.

XX Sequence 7379 BP; 2097 A; 1810 C; 1660 G; 1812 T; 0 other;

Query Match 76.3%; Score 20.6; DB 19; Length 7379;
 Best Local Similarity 85.2%; Pred. No. 71;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcagctcgacgtcgagggggg 27

Db 4299 GGGGCGCGGTGACCTCGAGGGGGG 4273

RESULT

ABL20258

ID ABL20258 standard; DNA; 23914 BP.

XX ABL20258;
 AC ABL20258;

XX 26-MAR-2002 (first entry)
 DT

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 12247.
 DE

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.
 OS

XX WO200171042-A2.
 PN

XX 27-SEP-2001.
 PD

XX 23-MAR-2001; 2001WO-US09231.
 PF

XX 23-MAR-2000; 2000US-191637P.
 PR

XX 11-JUL-2000; 2000US-0614150.
 PR

XX (PEKE) PE CORP NY.
 PA

XX Venter JC, Adams M, Li PWD, Myers EW;
 PI

XX WPI; 2001-656860/75.
 DR

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Claim 1; SEQ ID NO 12247; 2lpp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 23914 BP; 6269 A; 5220 C; 5706 G; 6719 T; 0 other;

Query Match 76.38; Score 20.6; DB 23914; Length 23914;
 Best Local Similarity 85.23; Pred. No. 64;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ggggtcgacgtcgacgtcgagggggg 27
 ||||| ||||| ||||| ||||| |||||
 Db 23139 ggcgtcgacgtcgacgtcgagggg 23165

RESULT 8
 AAS45342
 ID AAS45342 standard; DNA; 7317 BP.
 AC AAS45342;
 XX 18-DEC-2001 (first entry)
 DT DT
 DE Chemically pretreated genomic DNA associated with cell cycle #24.
 XX Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;
 KW graft-versus-host disease; glomerular disease; Lewy body disease; cancer;
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
 KW PCR primer.
 XX Homo sapiens.
 OS
 XX WO200168911-A2.
 PN 20-SEP-2001.
 PD
 XX 15-MAR-2001; 2001WO-EP02945.
 PF
 XX 15-MAR-2000; 2000DE-1013847.
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX (EPIG-) EPIGENOMICS AG.
 PA Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2001-602751/68.
 DR
 XX Designing primers and probes for analysing diseases associated with
 PT cytosine methylation state e.g. arthritis, cancer, aging,
 PT arteriosclerosis comprising fragments of chemically modified genes
 PT associated with cell cycle -
 XX
 PS Claim 1; SEQ ID No. 47; 28pp; English.
 CC Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA
 CC molecules associated with the cell cycle and specific PCR primers of the
 CC invention. The sequences are useful for detecting the methylation state
 CC of all CpG dinucleotides in a sequence and therefore for analysing
 CC associated diseases. By analysing cytosine methylations in the pretreated
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
 CC of existing diseases or the predisposition to specific diseases can be
 CC ascertained. The parameters may be compared to another set of genetic
 CC and/or epigenetic parameters, the differences serving as basis for
 CC diagnosis and/or prognosis events which are disadvantageous to patients.
 CC The sequences of the invention are useful for the diagnosis and therapy
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,
 CC aging, glomerular disease, Lewy body disease, arthritis,
 CC arteriosclerosis, solid tumours and cancers.
 XX Sequence 7317 BP; 1665 A; 326 C; 1881 G; 3445 T; 0 other;

Query Match 74.8%; Score 20.2; DB 22; Length 7317;
 Best Local Similarity 88.0%; Pred. No. 98;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ggggtcgacgtcgacgtcgagggggg 26
 ||||| ||||| ||||| ||||| |||||
 Db 5233 ggggtcgacgtcgacgttttagggggg 5257

RESULT 9
 AAV26974/c
 ID AAV26974 standard; RNA; 2032 BP.
 XX
 AC AAV26974;
 XX 11-SEP-1998 (first entry)
 DT DT
 DE Consensus I3L promoted FIV gag/protease expression cassette in vCP253.
 XX
 KW gag; protease; FIV; vaccinia; promoter; lentivirus; vaccine;
 KW feline immunodeficiency virus; ss.
 XX Vaccinia virus.
 OS Feline immunodeficiency virus.
 XX
 FH Key Location/Qualifiers
 FT Promoter .135..234
 FT /*tag= a
 FT /note= "Vaccinia I3L promoter"
 FT CDS 235..1648
 FT /*tag= b
 FT /note= "FIV gag/protease"
 XX
 PN WO9821354-A1.
 XX 22-MAY-1998.
 PD
 XX 07-NOV-1997; 97WO-US20430.
 PF
 XX 14-NOV-1996; 96US-0746668.
 PR
 XX (VIRO-) VIROGENETICS CORP.
 PA Paoletti E, Tartaglia J;
 XX WPI; 1998-297957/26.
 DR
 XX Vector containing exogenous DNA encoding a lentivirus epitope and
 PT derived expression products - or antibodies, useful in human or
 PT animal vaccines particularly against immune deficiency viruses
 XX
 PS Example 2; Fig 6; 123pp; English.
 CC The sequence is that of the vCP253 insertion, comprising the vaccinia
 CC I3L promoter and the FIV gag/protease ORF, which was used in the
 CC construction of a vector comprising comprises exogenous DNA (I)
 CC encoding at least one lentivirus epitope (II). Such a vector is
 CC useful in vaccines to generate a therapeutic or protective response
 CC in humans or animals against lentivirus, optionally in association
 CC with subsequent booster doses (of (II) or inactivated lentivirus).
 CC A particular application is protection against FIV infection. The
 CC vector is also used to express the products of (I) in vitro or
 CC ex vivo (i.e. with return of transfected cells to the patient).
 CC Antibodies may also be used for treatment or prevention of infection,
 CC or for diagnostic detection of lentiviral antigen or antibody. (I) is
 CC useful as a source of primers and hybridisation probes. Priming with
 CC the vector then boosting with corresponding subunit vaccine will
 CC protect against both homologous and heterologous strains. Since
 CC the vector lacks some virus-encoding genes, its virulence is
 CC attenuated and they are safer to use.
 CC

```
SQ Sequence 2032 BP; 747 A; 311 C; 435 G; 539 T; 0 other;

Query Match      73.3%; Score 19.8; DB 19; Length 2032;
Best Local Similarity 91.3%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 tcgacgtcgacgtcgagggggg 27
Db 133 TCGATGTCGACCTCGAGGGGGG 111

RESULT 10
AAV26976/c
ID AAV26976 standard; RNA; 4150 BP.
AC AAV26976;
XX
XX
XX
DT 11-SEP-1998 (first entry)
XX
XX
DE Consensus VCP329 H6 FIV 97TM/I3L FIV gag/pol expression cassette.
KW gag; pol; FIV; vaccinia; promoter; lentivirus; vaccine;
KW feline immunodeficiency virus; gp97; transmembrane domain; ss.
XX
XX
OS Vaccinia virus.
OS Feline immunodeficiency virus.
XX
XX
FH Key Location/Qualifiers
FT promoter complement(2146..2023)
FT /*tag= a
FT /note= "Vaccinia H6 promoter"
FT CDS complement(2022..42)
FT /*tag= b
FT /note= "FIV 97TM"
FT promoter 2253..2352
FT /*tag= a
FT /note= "Vaccinia I3L promoter"
FT CDS 2353..3766
FT /*tag= b
FT /note= "FIV gag/protease"
XX
XX
XX WO9821354-A1.
XX
XX 22-MAY-1998.
XX
XX 07-NOV-1997; 97WO-US20430.
XX
XX 14-NOV-1996; 96US-0746668.
XX
XX (VIRO-) VIROGENETICS CORP.
XX
XX Paolletti E, Tartaglia J;
XX
XX WPI: 1998-297957/26.
XX
XX
XX Vector containing exogenous DNA encoding a lentivirus epitope and
XX derived expression products - or antibodies, useful in human or
XX animal vaccines particularly against immune deficiency viruses
XX
XX Example 5; Fig 8; 123pp; English.
XX
XX The sequence is that of the vCP329 insertion, comprising the
XX vaccinia H6 promoted FIV 97TM/I3L promoted FIV gag/pol
XX expression cassette which was used in the
XX construction of a vector comprising comprises exogenous DNA (I)
XX encoding at least one lentivirus epitope (II). Such a vector is
XX useful in vaccines to generate a therapeutic or protective response
XX in humans or animals against lentivirus, optionally in association
XX with subsequent booster doses (of (II) or inactivated lentivirus).
XX A particular application is protection against FIV infection. The
XX vector is also used to express the products of (I) in vitro or
XX ex vivo (i.e. with return of transfected cells to the patient).
```

```
CC Antibodies may also be used for treatment or prevention of infection,
CC or for diagnostic detection of lentiviral antigen or antibody. (I) is
CC useful as a source of primers and hybridisation probes. Priming with
CC the vector then boosting with corresponding subunit vaccine will
CC protect against both homologous and heterologous strains. Since
CC the vector lacks some virus-encoding genes, its virulence is
CC attenuated and they are safer to use.
XX
XX Sequence 4150 BP; 1327 A; 771 C; 725 G; 1327 T; 0 other;

Query Match      73.3%; Score 19.8; DB 19; Length 4150;
Best Local Similarity 91.3%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 tcgacgtcgacgtcgagggggg 27
Db 2251 TCGATGTCGACCTCGAGGGGGG 2229

RESULT 11
AAA27756/c
ID AAA27756 standard; DNA; 1293 BP.
XX
XX AAA27756;
XX
XX 29-AUG-2000 (first entry)
XX
XX Neisseria meningitidis htrB1 gene region.
XX
XX Lipopolysaccharide; vaccine; adjuvant; htrB1 gene;
XX acyltransferase; toxicity; attenuation; ds.
XX
XX Neisseria meningitidis.
XX
XX Key Location/Qualifiers
FH CDS 1..331
FT /*tag= a
FT /partial
FT /codon_start= 2
FT /note= "ruvC gene, encodes AAY79682"
FT CDS 370..1251
FT /*tag= b
FT /note= "htrB1 gene, encodes AAY79683"
XX
XX WO200026384-A1.
XX
XX 11-MAY-2000.
XX
XX 03-NOV-1998; 98WO-NL00633.
XX
XX 03-NOV-1998; 98WO-NL00633.
XX
XX (NEWE-) NEDERLANDEN MIN WELZIJN.
XX
XX Van Der Ley PA, Hamstra HJ, Steeghs LJJM;
XX
XX WPI: 2000-422514/36.
XX
XX P-PSDB; AAY79682, AAY79683.
XX
XX New recombinant lipopolysaccharide, useful as low-toxicity adjuvant for
XX vaccines, has altered pattern of acylation and/or phosphate residues
XX attached to glucosamine.
XX
XX Disclosure; Fig 2B; 40pp; English.
XX
XX The present sequence is that of a chromosomal fragment of
XX Neisseria meningitidis DNA including the htrB1 gene. A BLAST
XX search on gonococcal genome sequences was made using htrB/msbB
XX gene sequences from Escherichia coli and Haemophilus influenzae.
XX Several contigs were identified, and PCR primers based on these
XX sequences were designed. The primers were used in the PCR
XX amplification of meningococcal chromosomal DNA to generate a 500
```

CC bp product. This was used as probe for the isolation of the
CC present, larger htrB1 gene. The htrB1 gene codes for an
CC acyltransferase (see AA79683). Involved in the secondary acylation
CC of lipopolysaccharide (LPS). Mutations in the htrB1 gene provide
CC an LPS product which is less toxic than native LPS, but which
CC has higher adjuvant activity. The invention is directed at novel
CC less toxic forms of LPS that are obtained through genetically
CC modified Gram-negative bacteria. The novel LPS has fewer secondary
CC acyl chains per molecule of LPS than the native LPS, the secondary
CC acyl chains being bound to primary acyl chains, and the primary
CC acyl chains being bound to the glucosamine of the LPS molecule.
CC Recombinant LPS is produced by cultivation of a Gram-negative
CC bacterium, such as *N. meningitidis*, having a mutation in a gene
CC encoding a protein involved in lipid A biosynthesis, particularly
CC at the level of secondary acyl addition, especially the htrB1 gene.
CC It is used as an adjuvant in vaccines used to stimulate an immune
CC response against Gram-negative bacteria, particularly for
CC controlling infections caused by organisms from which the LPS is
CC derived, or by other organisms. The acylation pattern of the
CC recombinant LPS is homogeneous, which facilitates standardization.
CC A partial open reading frame with homology to the *E. coli* rucC
CC gene is located upstream of the htrB1 gene in the present sequence.
XX
SQ Sequence 1293 BP; 314 A; 330 C; 346 G; 303 T; 0 other;

Query Match 72.6%; Score 19.6; DB 21; Length 1293;
Best Local Similarity 84.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 gggtcgacgtcgacgtcgagggggg 27
||| ||||| ||| ||||| |||||
Db 30 GGATTGACGTTCGCTCGAGGGGGG 5

RESULT 12
AAK52340
ID AAK52340 standard; cDNA; 2136 BP.
XX
AC AAK52340;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 885.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX

DR WPI; 2001-476283/51.
XX P-PSDB; AAM79207.

XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -

XX Claim 1; Page 2922-2924; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX
SQ Sequence 2136 BP; 683 A; 437 C; 512 G; 504 T; 0 other;

Query Match 72.6%; Score 19.6; DB 22; Length 2136;
Best Local Similarity 84.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 gggtcgacgtcgacgtcgagggggg 27
||| ||||| ||| ||||| |||||
Db 26 gggtcgacgatttcgagggggg 51

RESULT 13
AAI99683/c
ID AAI99683 standard; DNA; 4403765 BP.

XX
AC AAI99683;

XX
DT 15-JAN-2002 (first entry)

XX
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.

XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.

XX
OS Mycobacterium tuberculosis.

XX
PN US6294328-B1.

XX
PD 25-SEP-2001.

XX
PF 24-JUN-1998; 98US-0103840.

XX
PR 24-JUN-1998; 98US-0103840.

XX
PA (GENO-) INST GENOMIC RES.

XX
Fleischmann RD, White OR, Fraser CM, Venter JC;

XX
WPI; 2001-647261/74.

XX
XX Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where *M. tuberculosis* strains CDC
PT 1551 and H37Rv differ -

XX
XX Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.

XX
XX The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,

CC Mycobacterium tuberculosis or related Mycobacterium by determining the.
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
CC
SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;

Query Match 71.1%; Score 19.2; DB 22; Length 4403765;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggggtcagctgcagctcgagggg 24
||| ||||| ||||| ||||| |||||
Db 4368553 GGTGCCGACGGCGACGTCGAGGGG 4368530

RESULT 14
ID AAI99682/c
AAI99682 standard; DNA; 4411529 BP.
XX
AC AAI99682;
XX
XX
DT 15-JAN-2002 (first entry)
DE
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
XX Mycobacterium tuberculosis.
OS
XX
XX US6294328-B1.
XX
XX
PD 25-SEP-2001.
XX
XX 24-JUN-1998; 98US-0103840.
XX
XX 24-JUN-1998; 98US-0103840.
XX
XX (GENO-) INST GENOMIC RES.
XX
XX Fleischmann RD, White OR, Fraser CM, Venter JC;
XX
XX WPI; 2001-647261/74.
XX
XX Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ -
XX
XX
PS Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
XX The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.

XX

SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;

Query Match 71.1%; Score 19.2; DB 22; Length 4411529;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggggtcagctgcagctcgagggg 24
||| ||||| ||||| ||||| |||||
Db 4376321 GGTGCCGACGGCGACGTCGAGGGG 4376298

RESULT 15

AAA31424

ID AAA31424 standard; DNA; 400 BP.

XX

AC AAA31424;

XX

DT 05-JUL-2000 (first entry)

DE

DE Plant microsatellite marker #385.

XX

KW Plant microsatellite sequence; core repeat sequence; detection; probe;
KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;
KW variety identification; genetic variability evaluation; primer; ss.

XX

OS Eucalyptus grandis.

XX

PN WO9967421-A1.

XX

PD 29-DEC-1999.

XX

PF 25-JUN-1999; 99WO-NZ00092.

XX

PR 25-JUN-1998; 98US-0105307.

XX

PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX

PI Havukkala IJ, Bloksberg LN, Glenn M;

XX

DR WPI; 2000-116958/10.

XX

PT New plant microsatellite markers and associated flanking species for
PT the detection of polymorphic genetic markers -

XX

PS Claim 1; Page 190; 392pp; English.

XX

CC Sequences AAA31040-A32093 represent novel plant microsatellite sequences
CC and associated flanking species. The sequences comprise a central core
CC repeat sequence, especially selected from the sequences AAA32094-A32096
CC with left and right flanking sequences. The polynucleotide sequences
CC can be used in the detection of DNA polymorphisms, in genome mapping,
CC in physical mapping, in positional cloning of genes, in variety
CC identification and in evaluation of genetic variability within and
CC between plant tissues, populations, cultivars, species and species
CC groups. They may also be used to design hybridization probes for
CC oligonucleotide fingerprinting and library screening and to design
CC primers for microsatellite-primed PCR. Microsatellite markers are
CC useful to locate specific economically useful genes in plant genomes.

XX

SQ Sequence 400 BP; 87 A; 149 C; 113 G; 51 T; 0 other;

Query Match

Best Local Similarity 70.4%; Score 19; DB 21; Length 400;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ggggtcagctgcagctcgagggggg 27

||||| ||||| ||| ||| ||||| |||

Db 362 ggggacagcggcgccgagggggag 388

Sat Aug 10 09:08:34 2002

Search completed: August 10, 2002, 03:24:27
Job time: 13838 sec

us-09-672-126-36.rng

Page 9

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:06:33 ; Search time 277.54 seconds
(without alignments)
23.896 Million cell updates/sec

Title: US-09-672-126-36
Sequence: 1 ggggtcgacgtcgacgtcgagggggg 27

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/6C_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20.6	76.3	1438	3	US-08-434-099A-26
C 2	19.2	76.3	7379	3	US-08-675-566-13
C 3	20.6	71.1	4403765	4	US-09-103-840A-2
C 4	19.2	71.1	4411529	4	US-09-103-840A-1
C 5	18.2	67.4	3306	1	US-08-261-206A-71
C 6	18.2	67.4	28958	1	US-08-258-261B-6
C 7	18.2	67.4	28958	1	US-08-456-837-6
C 8	18.2	67.4	28958	1	US-08-457-342-6
C 9	18.2	67.4	28958	1	US-08-457-646A-6
C 10	18.2	67.4	28958	1	US-08-458-076A-6
C 11	18.2	67.4	28958	1	US-08-764-233A-4
C 12	18.2	67.4	28958	1	US-08-457-335A-6
C 13	18.2	67.4	28958	1	US-08-729-214-6
C 14	18.2	67.4	28958	3	US-09-028-934-6
C 15	18.2	67.4	49377	1	US-08-764-233A-1
C 16	18	66.7	1342	4	US-09-500-569-9
C 17	18	66.7	68750	3	US-09-335-409-1
C 18	18	66.7	68750	4	US-09-568-102-1
C 19	18	66.7	68750	4	US-09-567-969-1
C 20	18	66.7	68750	4	US-09-568-480-1
C 21	18	66.7	68750	4	US-09-568-486-1
C 22	18	66.7	68750	4	US-09-568-472-1
C 23	17.8	65.9	474	3	US-08-928-799A-2
C 24	17.8	65.9	898	2	US-08-997-080-185
C 25	17.8	65.9	898	2	US-08-997-362-185
C 26	17.8	65.9	898	2	US-08-995-855-185
C 27	17.8	65.9	898	4	US-09-324-542-185

28	17.8	65.9	1364	4	US-09-095-855-204
29	17.6	65.2	40	1	US-08-244-378A-27
30	17.6	65.2	503	3	US-08-581-918A-35
31	17.6	65.2	503	4	US-08-346-147B-35
32	17.6	65.2	515	3	US-08-581-918A-36
33	17.6	65.2	515	4	US-08-346-147B-36
34	17.6	65.2	2646	1	US-08-539-304A-5
35	17.6	65.2	4403765	4	US-09-103-840A-2
36	17.6	65.2	4411529	4	US-09-103-840A-1
C 37	17.4	64.4	28	1	US-08-324-001-23
C 38	17.4	64.4	31	2	US-08-553-339-5
C 39	17.4	64.4	31	2	US-09-061-542-5
C 40	17.4	64.4	31	4	US-08-450-274-5
C 41	17.4	64.4	31	5	PCT-US94-05285A-5
C 42	17.4	64.4	42	4	US-09-180-143-1
C 43	17.4	64.4	57	4	US-09-203-681-4
44	17.4	64.4	70	1	US-08-144-602B-20
45	17.4	64.4	88	1	US-08-144-602B-15

ALIGNMENTS

RESULT 1
US-08-434-099A-26/c
; Sequence 26, Application US/08434099A
; Patent No. 6083902
; GENERAL INFORMATION:
; APPLICANT: Cederholm-Mms., Stewart A.
; TITLE OF INVENTION: Recombinant Fibrin Chains,
; TITLE OF INVENTION: Fibrin and Fibrin-Homologs
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.R. Squibb & Sons, Inc.
; STREET: 100 Headquarters Park Drive
; CITY: Skillman
; STATE: NJ
; COUNTRY: USA
; ZIP: 08558
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,099A
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/236,979
; FILING DATE: 02-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Furman, Jr., Esq., Theodore R
; REGISTRATION NUMBER: 30,942
; REFERENCE/DOCKET NUMBER: CV0054a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-281-2372
; TELEFAX: 908-281-2373
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1438 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 3...1364
; OTHER INFORMATION:
US-08-434-099A-26

Query Match 76.3% ; Score 20.6 ; DB 3 ; Length 1438 ;

Best Local Similarity 85.2%; Pred. No. 20;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcgcacgtcgacgtcgagggggg 27
||||| ||||| ||||| ||||| |||||
Db 36 GGGCCCGCGTCGACCTCGAGGGGGG 10

RESULT 2
US-08-675-566-13/c
; Sequence 13, Application US/08675566
; Patent No. 6090393
; GENERAL INFORMATION:
; APPLICANT: Fischer, Laurent
; TITLE OF INVENTION: PROMOTERS, EXPRESSION CASSETTES,
; TITLE OF INVENTION: RECOMBINANT VIRUSES, METHODS FOR MAKING, AND USES THEREOF
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,566
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2890
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)840-3333
; TELEFAX: (212)840-0712
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7379 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-675-566-13

Query Match 76.3%; Score 20.6; DB 3; Length 7379;
Best Local Similarity 85.2%; Pred. No. 17;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggggtcgcacgtcgacgtcgagggggg 27
||||| ||||| ||||| ||||| |||||
Db 4299 GGGCCCGCGTCGACCTCGAGGGGGG 4273

RESULT 3
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24

; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 71.1%; Score 19.2; DB 4; Length 4403765;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggggtcgcacgtcgacgtcgagggg 24
||||| ||||| ||||| ||||| |||||
Db 4368553 GGTGCCGACGGCGACGTGAGGGG 4368530

RESULT 4
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 71.1%; Score 19.2; DB 4; Length 4411529;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggggtcgcacgtcgacgtcgagggg 24
||||| ||||| ||||| ||||| |||||
Db 4376321 GGTGCCGACGGCGACGTGAGGGG 4376298

RESULT 5
US-08-261-206A-71/c
; Sequence 71, Application US/08261206A
; Patent No. 5574007
; GENERAL INFORMATION:
; APPLICANT: Zushi, Mitichitaka
; APPLICANT: Gomi, Komakazu
; APPLICANT: Yamamoto, Shuji
; APPLICANT: Suzuki, Koji
; APPLICANT: Matsuda, Akio
; TITLE OF INVENTION: A Polypeptide Capable of Interacting
; TITLE OF INVENTION: with Thrombin
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N Washington St.
; CITY: Falls Church
; STATE: Virginia

;; COUNTRY: USA
;; ZIP: 22046-0747
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/261.206A
;; FILING DATE:
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/740.492
;; FILING DATE: 03-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Svensson, Leonard R.
;; REGISTRATION NUMBER: 30330
;; REFERENCE/DOCKET NUMBER: 216-275P
;; TELEPHONE: 703-241-1300
;; TELEFAX: 703-241-2848
;; TELEX: 248345
;; INFORMATION FOR SEQ ID NO: 71:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3306 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; ORIGINAL SOURCE:
;; ORGANISM: Acremonium chrysogenum
;; FEATURE:
;; NAME/KEY: -
;; LOCATION: 1..3306
;; OTHER INFORMATION: /label= PKG_gene
;; OTHER INFORMATION: /note= "Nucleotide sequence of region A in Figure
;; OTHER INFORMATION: 59. The sequence is presented as Figure 61."
;; FEATURE:
;; NAME/KEY: exon
;; LOCATION: 1252..1317
;; FEATURE:
;; NAME/KEY: exon
;; LOCATION: 1463..1883
;; FEATURE:
;; NAME/KEY: exon
;; LOCATION: 1948..2715
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: Join(1252..1317, 1463..1883, 1948..2714)
US-08-261-206A-71

Query Match 67.4%; Score 18.2; DB 1; Length 3306;
Best Local Similarity 87.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 gtcgacgtcgacgtcgagggggg 26
DB 2543 GTCGACGACGCGTCGAGGGTGG 2521

RESULT 6
US-08-258-261B-6/c
; Sequence 6, Application US/08258261B
; Patent No. 5639949
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting

;; APPLICANT: Hammer, Phillip E.
;; APPLICANT: Uknes, Scott Joseph
;; TITLE OF INVENTION: Genes for the synthesis of
;; TITLE OF INVENTION: antipathogenic substances
;; NUMBER OF SEQUENCES: 22
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ciba-Geigy Corporation
;; STREET: 7 Skyline Drive
;; CITY: Hawthorne
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10532
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/258.261B
;; FILING DATE: 08-JUN-1994
;; CLASSIFICATION: 800
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/457.205
;; FILING DATE: 01-JUN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Elmer, James Scott
;; REGISTRATION NUMBER: 36,129
;; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8614
;; TELEFAX: 919-541-8689
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 28958 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-258-261B-6

Query Match 67.4%; Score 18.2; DB 1; Length 28958;
Best Local Similarity 87.0%; Pred. No. 93;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 gtcgacgtcgacgtcgagggggg 26
DB 24795 GACGACGTCGACGTCGGCGGGG 24773

RESULT 7
US-08-456-837-6/c
; Sequence 6, Application US/08456837
; Patent No. 5643774
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
;; TITLE OF INVENTION: Genes for the synthesis of
;; TITLE OF INVENTION: antipathogenic substances
;; NUMBER OF SEQUENCES: 22
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ciba-Geigy Corporation
;; STREET: 7 Skyline Drive
;; CITY: Hawthorne

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,342
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 424
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
; PS-08-457-342-6

```

Gaps 0;

Oy 4 atcgaatcgacgtcgaagggg 26

Db 24795 GACGACGTCGACGTCGGCGGGG 24773

US-08-437-646A-6/C
: Sequence 6. Application US/08457646A

APPLICANT: Ligon, James M.
 APPLICANT: Beck, James Joseph
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Rvals, John Andrew
 APPLICANT: Gaffney, Thomas Deane
 APPLICANT: Lam, Stephen Ting
 APPLICANT: Hammer, Phillip E.
 APPLICANT: Uknes, Scott Joseph
 TITLE OF INVENTION: Genes for the synthesis of
 antiproliferative substances
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/457,646A
 FILING DATE: 01-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/457,205

```

; NAME: Elmer, James Scott
;
; REGISTRATION NUMBER: 36,129
;
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3

```

TOPOLOGY: linear

```

: NAME: Elmer, James Scott
:
: REGISTRATION NUMBER: 36,129
:
: REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
:

```

RESULT 13

US 00 723 214 0/C
; Sequence 6, Application US/08729214
; Patent No. 5817502

```

: GENERAL INFORMATION:
:
: APPLICANT: Ligon, James M.
: APPLICANT: Hill, Dwight Steven
: APPLICANT: Ryals, John Andrew
: APPLICANT: Hammer, Phillip E.
: APPLICANT: van Pee, Karl-Heinz
: APPLICANT: Kirner, Sabine
:
: TITLE OF INVENTION: Genes for t
:
: TITLE OF INVENTION: antipathoge
:
: NUMBER OF SEQUENCES: 27
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Ciba-Geigy Corpora
:
: STREET: 520 White Plains Road
:
: CITY: Tarrytown
:
: STATE: NY
:
: COUNTRY: USA
:
: ZIP: 10591

```

```

; ; COMPUTER READABLE FORM:
; ; MEDIUM TYPE: Floppy disk
; ; COMPUTER: IBM PC compatible
; ; OPERATING SYSTEM: PC-DOS/MS-DOS
; ; SOFTWARE: Patent In Release #1.0, Version #1.25
; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: US/08/729,214
; ; FILING DATE: TBA
; ; CLASSIFICATION: 435
; ; ATTORNEY/AGENT INFORMATION:

```

NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1506/CIP5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-729-214-6

Query Match 67.4%; Score 18.2; DB 1; Length 28958;
Best Local Similarity 87.0%; Pred. No. 93;
Matches 20: Conservative 0; Mismatches 3; Indels 0;

Qy 4 gtcgacgtcgacgtcgagggggg 26

Qy 4 gtcgacgtcgacgtcgagggggg 26
| | | | |
Db 24795 GACGACGTTCGACGTCTCGGGGGG 24773

RESULT 14
US-09-028-934-6/c
; Sequence 6, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: LIGON, James M.

APPLICANT: Ligon, James M.


```
; OTHER INFORMATION: /product= "Module 1 of SorB"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 24638..30820
; OTHER INFORMATION: /product= "Module 2 of SorB"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 30881..35446
; OTHER INFORMATION: /product= "Module 3 of SorB"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 35528..40114
; OTHER INFORMATION: /product= "Module 4 of SorB"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 40190..46318
; OTHER INFORMATION: /product= "Module 5 of SorB"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 46851..47891
; OTHER INFORMATION: /product= "Sorm"
; OTHER INFORMATION: /note= "The protein encoded by the sorm gene is highly
; OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
; OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
; OTHER INFORMATION: polyketide rappamycin."
; US-08-764-233A-1
```

```
Query Match 67.4%; Score 18.2; DB 1; Length 49377;
Best Local Similarity 87.0%; Pred. No. 87;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 gtctgacgtcgacgtcgagggggg 26
| | | | | | | | | | | | | | | |
Db 41098 GACGACGTCGACGTCGCGGGG 41076
```

Search completed: August 10, 2002, 03:09:20
Job time: 16226 sec

Best Local Similarity 85.2%; Pred. No. 4.2e+03;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:58:57 ; Search time 2778.35 seconds
(without alignments)
158.172 Million cell updates/sec

Title: US-09-672-126-37
Perfect score: 21
Sequence: 1 gggagcagctgctggggggg 21

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_on.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	21	100.0	21	6	AX104887	Sequence
2	21	100.0	21	6	AX105139	Sequence
3	19.4	92.4	22	6	AX104797	Sequence
4	19.4	92.4	22	6	AX105112	Sequence
5	19.4	92.4	118905	2	AC096093	Sequence
6	18.4	87.6	20	6	AX104884	Sequence
7	18.4	87.6	20	6	AX105262	Sequence
8	17.8	84.8	22	6	AX104848	Sequence
9	17.8	84.8	22	6	AX105123	Sequence
10	17.8	84.8	30853	2	AC094245	Sequence
11	17.8	84.8	61633	2	AC084075	Sequence
12	17.8	84.8	136120	8	AC068923	Sequence
13	17.8	84.8	157246	2	AC095209	Sequence
14	17.8	84.8	160867	2	AC105838	Sequence
15	17.8	84.8	166036	2	AC094985	Sequence
16	17.8	84.8	184427	14	EHVU20824	Sequence
17	17.8	84.8	195917	2	AC091289	Sequence
18	17.4	82.9	39726	1	SCBD11	Sequence
19	16.8	80.0	20	6	AX104717	Sequence
20	16.8	80.0	20	6	AX104780	Sequence
21	16.8	80.0	20	6	AX104882	Sequence
22	16.8	80.0	20	6	AX105107	Sequence
23	16.8	80.0	20	6	AX105237	Sequence
24	16.8	80.0	20	6	AX105253	Sequence
25	16.8	80.0	20	6	AX105261	Sequence
26	16.8	80.0	20	6	AX355415	Sequence
27	16.8	80.0	2224	3	AF173262	Sequence
28	16.8	80.0	36307	9	HS366D1	Sequence
29	16.8	80.0	74265	8	AC068901	Sequence
30	16.8	80.0	74881	2	AC020320	Sequence
31	16.8	80.0	94908	2	AC099247	Sequence
32	16.8	80.0	114285	2	AC094938	Sequence
33	16.8	80.0	119915	2	AC103240	Sequence
34	16.8	80.0	122349	1	D90908	Sequence
35	16.8	80.0	148407	2	AC097603	Sequence
36	16.8	80.0	158810	2	AC095530	Sequence
37	16.8	80.0	160480	8	AF123535	Sequence
38	16.8	80.0	166050	1	AL646085	Sequence
39	16.8	80.0	166510	2	AC098703	Sequence
40	16.8	80.0	171831	3	AC007473	Sequence
41	16.8	80.0	278196	3	AE003825	Sequence
42	16.4	78.1	469	5	D88114	Sequence
43	16.4	78.1	556	5	SSHEMA	Sequence
44	16.4	78.1	175947	8	AP003435	Sequence
45	16.4	78.1	209150	2	AC023642	Sequence

ALIGNMENTS

RESULT 1	AX104887	Sequence 1079 from Patent WO0122972.	21 bp	DNA	linear	PAT 30-APR-2001
LOCUS	AX104887	Sequence 1079 from Patent WO0122972.				
DEFINITION	AX104887	Sequence 1079 from Patent WO0122972.				
ACCESSION	AX104887	Sequence 1079 from Patent WO0122972.				
VERSION	AX104887.1	GI:13921084				
KEYWORDS		synthetic construct.				
SOURCE		synthetic construct.				
ORGANISM		artificial sequence.				
REFERENCE		1 (bases 1 to 21)				
AUTHORS		Krieg, A.M., Schetter, C. and Vollmer, J.C.				
TITLE		Immunostimulatory nucleic acids				
JOURNAL		Patent: WO 0122972-A 1079 05-APR-2001;				
		UNIVERSITY OF IOWA RESEARCH FOUNDATION, (US) ; Coley Pharmaceutical				
		GmbH (DE)				
FEATURES		Location/Qualifiers				
Source		1..21				
		/organism="synthetic construct"				
		/db_xref="taxon:32630"				
BASE COUNT		2 a 3 c 14 g 2 t				
ORIGIN						

```

Query Match      100.0%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgacgtcgtggggggg 21
Db 1 GGGGACGACGTCGTCGGGGGGG 21

RESULT 2
AX105139
LOCUS AX105139 21 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 37 from Patent WO0122990.
ACCESSION AX105139
VERSION AX105139.1 GI:13921289
KEYWORDS
SOURCE
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 21)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 37 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source
1..21
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
misc_feature
1..2
/note="Backbone has phosphorothioate linkages."
misc_feature
3..15
/note="Backbone has phosphodiester linkages."
misc_feature
16..20
/note="Backbone has phosphorothioate linkages."
misc_feature
21
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 3 c 14 g 2 t
ORIGIN
1
ggggacgacgtcgtggggggg 21
GGGGACGACGTCGTCGGGGGGG 21

Query Match      100.0%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgacgtcgtggggggg 21
Db 1 GGGGACGACGTCGTCGGGGGGG 21

RESULT 3
AX104797
LOCUS AX104797 22 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 989 from Patent WO0122972.
ACCESSION AX104797
VERSION AX104797.1 GI:13920994
KEYWORDS
SOURCE
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 989 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES
source
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"

```

```

BASE COUNT 2 a 4 c 14 g 2 t

Query Match      92.4%; Score 19.4; DB 6; Length 22;
Best Local Similarity 95.2%; Pred. No. 4.5e+03;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgacgtcgtggggggg 21
Db 2 GGGGACGACGTCGTCGGGGGGG 22

RESULT 4
AX105112
LOCUS AX105112 22 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 10 from Patent WO0122990.
ACCESSION AX105112
VERSION AX105112.1 GI:13921262
KEYWORDS
SOURCE
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 10 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
FEATURES
source
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide"
misc_feature
1..2
/note="Backbone has phosphorothioate linkages."
misc_feature
3..16
/note="Backbone has phosphodiester linkages."
misc_feature
17..21
/note="Backbone has phosphorothioate linkages."
misc_feature
22
/note="Backbone has phosphodiester linkages."
BASE COUNT 2 a 4 c 14 g 2 t
ORIGIN
1
ggggacgacgtcgtggggggg 21
GGGGACGACGTCGTCGGGGGGG 22

Query Match      92.4%; Score 19.4; DB 6; Length 22;
Best Local Similarity 95.2%; Pred. No. 4.5e+03;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgacgtcgtggggggg 21
Db 2 GGGGACGACGTCGTCGGGGGGG 22

RESULT 5
AC096093/c
LOCUS AC096093/c 118905 bp DNA linear HTG 20-DEC-2001
DEFINITION Rattus norvegicus clone CH230-24f4, *** SEQUENCING IN PROGRESS ***,
56 unordered pieces.
ACCESSION AC096093
VERSION AC096093.3 GI:17943776
KEYWORDS HTG; HTGS-PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 118905)
AUTHORS Muzny,D.M., Adams,C., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,

```

Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chondry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, N.D., Dathorne, S.K., David, K., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, B., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H., Lozano, R.J., Lu, X., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

Unpublished
2 (bases 1 to 118905)
Worley, K.C.

Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:16901731.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GBKO
Center clone name: CH230-24F4
----- Summary Statistics -----
Assembly program: Phrap; version 0.990329First call to findPhrapList
Consensus quality: 91369 bases at least Q40
Consensus quality: 101050 bases at least Q30
Consensus quality: 108564 bases at least Q20
Estimated insert size: 85142; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 0.9x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 56 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
4470
4570
8063
8163
11706
11806
15576
15676
18499
18599
21357
21456
24276
24376
27684
27784
30229
30530
33624
33625
33725
35946
37046
38755
38854
41454
41554
41555
44846
44946
47832
47932
49895
49896
52106
52206
54304
54305
57586
57587
59360
59460
61112
61113
62123
62452
62553
63709
63710
63810
65474
65574
68302
68303
68402
70830
70929
72688
72689
72789
74851
74951
76511
76512
78605
80752
80852
80853
82161
82261
84332
84333
85788
85789
85889
87524
87525
89192
89193
90667

contig of 4469 bp in length
gap of unknown length
contig of 3493 bp in length
gap of unknown length
contig of 3543 bp in length
gap of unknown length
contig of 3770 bp in length
gap of unknown length
contig of 2824 bp in length
gap of unknown length
contig of 2757 bp in length
gap of unknown length
contig of 2820 bp in length
gap of unknown length
contig of 3308 bp in length
gap of unknown length
contig of 2745 bp in length
gap of unknown length
contig of 2995 bp in length
gap of unknown length
contig of 3221 bp in length
gap of unknown length
contig of 1709 bp in length
gap of unknown length
contig of 2600 bp in length
gap of unknown length
contig of 3291 bp in length
gap of unknown length
contig of 2111 bp in length
gap of unknown length
contig of 2098 bp in length
gap of unknown length
contig of 3182 bp in length
gap of unknown length
contig of 1674 bp in length
gap of unknown length
contig of 1652 bp in length
gap of unknown length
contig of 1240 bp in length
gap of unknown length
contig of 1157 bp in length
gap of unknown length
contig of 1664 bp in length
gap of unknown length
contig of 2729 bp in length
gap of unknown length
contig of 2427 bp in length
gap of unknown length
contig of 1759 bp in length
gap of unknown length
contig of 2062 bp in length
gap of unknown length
contig of 1561 bp in length
gap of unknown length
contig of 1893 bp in length
gap of unknown length
contig of 2148 bp in length
gap of unknown length
contig of 1308 bp in length
gap of unknown length
contig of 2072 bp in length
gap of unknown length
contig of 1356 bp in length
gap of unknown length
contig of 1636 bp in length
gap of unknown length
contig of 1568 bp in length
gap of unknown length
contig of 1375 bp in length

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL

COMMENT

* 90668 90767: gap of unknown length.
* 90768 92132: contig of 1365 bp in length
* 92133 92232: gap of unknown length
* 92233 93014: contig of 1682 bp in length
* 93015 93915: gap of unknown length
* 93916 94015: contig of 1596 bp in length
* 94016 95610: gap of unknown length
* 95611 95710: contig of 1237 bp in length
* 95711 96947: gap of unknown length
* 96948 97048: contig of 1186 bp in length
* 97049 98233: gap of unknown length
* 98234 98334: gap of unknown length
* 98335 100151: contig of 1818 bp in length
* 100152 100231: gap of unknown length
* 100232 102147: contig of 1896 bp in length
* 102148 102247: gap of unknown length
* 102249 103874: contig of 1627 bp in length
* 103875 103974: gap of unknown length
* 103975 105273: contig of 1299 bp in length
* 105274 105373: gap of unknown length
* 105374 106386: contig of 1013 bp in length
* 106387 106487: gap of unknown length
* 106488 107729: contig of 1243 bp in length
* 107730 107730: gap of unknown length
* 107731 109139: contig of 1310 bp in length
* 109140 109239: gap of unknown length
* 109240 110456: contig of 1217 bp in length
* 110457 110556: gap of unknown length
* 110557 111603: contig of 1047 bp in length
* 111604 111703: gap of unknown length
* 111704 113460: contig of 1757 bp in length
* 113461 113560: gap of unknown length
* 113561 114631: contig of 1071 bp in length
* 114632 114731: gap of unknown length
* 114732 116226: contig of 1495 bp in length
* 116227 116327: gap of unknown length
* 116328 117472: contig of 1146 bp in length
* 117473 117572: gap of unknown length
* 117573 118905: contig of 1333 bp in length.
Location/Qualifiers
Query Match 92.4%; Score 19.4; DB 2; Length 118905;
Best Local Similarity 95.2%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgacgtcggtggggg 21
||||| ||||||| ||||||| |||||||
Db 37841 GGGCCGACGTCGTCGGGGG 37821

RESULT 6
AX104884
LOCUS AX104884 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1076 from Patent WO0122972.
ACCESSION AX104884
VERSION AX104884.1 GI:13921081
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Krieg, A.M., Schetter, C. and Vollmer, J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 1076 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US); Coley Pharmaceutical
GmbH (DE)
Location/Qualifiers
FEATURES
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 1 a 3 c 13 g 3 t
ORIGIN

QY 1 ggggacgacgtcggtggggg 20
||||| ||||||| ||||||| |||||||
Db 1 GGGGACGTCGTCGGGGG 20

Query Match 87.6%; Score 18.4; DB 6; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.1e+04;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ggggacgacgtcggtggggg 20
||||| ||||||| ||||||| |||||||
Db 1 GGGGACGTCGTCGGGGG 20
RESULT 7
AX105262
LOCUS AX105262 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 161 from Patent WO0122990.
ACCESSION AX105262
VERSION AX105262.1 GI:13921412
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Hartmann, G.D., Bratzler, R.L., and Krieg, A.U.
TITLE Methods related to immunostimulatory nucleic acid-induced
interferon
JOURNAL Patent: WO 0122990-A 161 05-APR-2001;
Coley Pharmaceutical Group, Inc. (US); UNIVERSITY OF IOWA RESEARCH
FOUNDATION (US)
Location/Qualifiers
FEATURES
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 1..2
/note="Synthetic Oligonucleotide"
misc_feature 3..14
/note="Backbone has phosphorothioate linkages."
misc_feature 15..19
/note="Backbone has phosphodiester linkages."
misc_feature 20
/note="Backbone has phosphorothioate linkages."
BASE COUNT 1 a 3 c 13 g 3 t
ORIGIN

Query Match 87.6%; Score 18.4; DB 6; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.1e+04;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ggggacgacgtcggtggggg 20
||||| ||||||| ||||||| |||||||
Db 1 GGGGACGTCGTCGGGGG 20

RESULT 8
AX104848
LOCUS AX104848 22 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1040 from Patent WO0122972.
ACCESSION AX104848
VERSION AX104848.1 GI:13921045
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS Krieg, A.M., Schetter, C. and Vollmer, J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 1040 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US); Coley Pharmaceutical
GmbH (DE)
Location/Qualifiers
FEATURES
source 1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 3 a 3 c 13 g 3 t
ORIGIN

ORIGIN

Query Match 84.8%; Score 17.8; DB 6; Length 22;
 Best Local Similarity 90.5%; Pred. No. 1.7e+04;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgcgtctgtgggggg 21
 ||| ||||| ||||| ||||| |||||
 Db 2 GGGACGAGCTGTTGGGGG 22

RESULT 9

AX105123 22 bp DNA linear PAT 30-APR-2001
 LOCUS Sequence 21 from Patent WO0122990.
 DEFINITION AX105123
 ACCESSION AX105123
 VERSION AX105123.1 GI:13921273
 KEYWORDS

SOURCE
 ORGANISM
 synthetic construct.
 synthetic construct.
 artificial sequence.

REFERENCE 1 (bases 1 to 22)

AUTHORS Hartmann,G.D., Bratler,R.L. and Krieg,A.U.

TITLE Methods related to immunostimulatory nucleic acid-induced

INTERFERON

JOURNAL Patent: WO 0122990-A 21 05-APR-2001;

Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)

FEATURES

source

Location/Qualifiers

1..22

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="Synthetic Oligonucleotide"

misc_feature

1..22 /note="Backbone has phosphorothioate linkages."

3..16

misc_feature

/note="Backbone has phosphodiester linkages."

17..21

misc_feature

/note="Backbone has phosphorothioate linkages."

22

misc_feature

/note="Backbone has phosphodiester linkages."

3 a 3 c 13 g 3 t

BASE COUNT

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

3 a 3 c 13 g 3 t

Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,C., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,X., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissegh,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,S., Nguyen,N., Nickerson,E., Nwokenwo,S., Ogutu,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D., Weinstein,G., and Gibbs,R.

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GAGP
 Center clone name: CH230-3E2
 Summary Statistics
 Assembly program: Phrap; version 0.990329First call to findPhrapList
 Consensus quality: 25262 bases at least Q40
 Consensus quality: 28277 bases at least Q30
 Consensus quality: 30421 bases at least Q20
 Estimated insert size: 13576; sum-of-contigs estimation
 Quality coverage: 0x in Q20 bases; agarose-fp estimation
 Quality coverage: 0.1x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 20 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2255: contig of 2255 bp in length

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Barna,N., Bastien,V., Bada,F., Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-pierre,N., Hagos,B., Heaford,A., Horton,L., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Triglio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE

Submitted (12-OCT-2000) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L10806

Center clone name: 2380_K_20

* NOTE: This record contains 79 individual

* sequencing reads that have not been assembled into

* contigs. Runs of N are used to separate the reads

* and the order in which they appear is completely

* arbitrary. Low-pass sequence sampling is useful for

* identifying clones that may be gene-rich and allows

* overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone

* will be sequenced to completion. In the event that

* the record is updated, the accession number will

* be preserved.

* 1

* 664 763: gap of 663 bp in length

* 764 1437: contig of 674 bp in length

* 1438 1537: gap of 100 bp

* 1538 2244: contig of 707 bp in length

* 2245 2344: gap of 100 bp

* 2345 3028: contig of 684 bp in length

* 3029 3128: gap of 100 bp

* 3129 3803: contig of 675 bp in length

* 3804 3903: gap of 100 bp

* 3904 4582: contig of 679 bp in length

* 4583 4682: gap of 100 bp

* 4683 5371: contig of 689 bp in length

* 5372 5471: gap of 100 bp

* 5472 6140: contig of 669 bp in length

* 6141 6240: gap of 100 bp

* 6241 6914: contig of 674 bp in length

* 6915 7014: gap of 100 bp

* 7015 7686: contig of 672 bp in length

* 7687 7786: gap of 100 bp

* 7787 8477: contig of 691 bp in length

* 8478 8577: gap of 100 bp

* 8578 9276: contig of 699 bp in length

* 9277 9376: gap of 100 bp

* 9377 10046: contig of 670 bp in length

* 10047 10146: gap of 100 bp

* 10147 10835: contig of 689 bp in length

* 2256

* 2355: gap of unknown length

* 2356

* 3730: contig of 1375 bp in length

* 3731

* 3830: gap of unknown length

* 3831

* 5870: contig of 2040 bp in length

* 5871

* 5970: gap of unknown length

* 5971

* 7795: contig of 1825 bp in length

* 7796

* 7895: gap of unknown length

* 7896

* 7895: gap of unknown length

* 9551

* 9650: gap of unknown length

* 9651

* 10735: contig of 1085 bp in length

* 10736

* 10835: gap of unknown length

* 12936

* 12994: contig of 1459 bp in length

* 12995

* 12994: gap of unknown length

* 13295

* 13988: contig of 1594 bp in length

* 13989

* 14088: gap of unknown length

* 14089

* 15363: contig of 1275 bp in length

* 15364

* 15463: gap of unknown length

* 17137

* 17136: contig of 1673 bp in length

* 17237

* 17236: gap of unknown length

* 17237

* 18725: contig of 1488 bp in length

* 18725

* 18824: gap of unknown length

* 18825

* 20353: contig of 1529 bp in length

* 20354

* 20453: gap of unknown length

* 20454

* 21457: contig of 1004 bp in length

* 21458

* 21557: gap of unknown length

* 21558

* 22708: contig of 1151 bp in length

* 22709

* 22808: gap of unknown length

* 22809

* 24338: contig of 1530 bp in length

* 24339

* 24438: gap of unknown length

* 24439

* 25457: contig of 1019 bp in length

* 25458

* 25557: gap of unknown length

* 25558

* 26880: contig of 1323 bp in length

* 26881

* 26980: gap of unknown length

* 26981

* 27999: contig of 1019 bp in length

* 28000

* 28099: gap of unknown length

* 28100

* 29704: contig of 1605 bp in length

* 29705

* 29804: gap of unknown length

* 29805

* 30853: contig of 1049 bp in length.

FEATURES

source

1. 30853

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

/clone="CH230-3E2"

7624 a 6656 c 6470 g 8108 t 1995 others

BASE COUNT

ORIGIN

Query Match 84.8%; Score 17.8; DB 2; Length 30853;

Best Local Similarity 90.5%; Pred. No. 1.7e+03;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggagcagctcgtgggggg 21

||||||| | |||||

Db 26350 GGGGACGAGGACGTGGGGGG 26330

RESULT 11

AC084075/c

LOCUS

AC084075 61633 bp DNA linear HTG 12-OCT-2000

DEFINITION

Homo sapiens chromosome 11 clone CTD-2380K20 map 11, LOW-PASS

AC084075

AC084075

VERSION

AC084075.1 GI:10799439

KEYWORDS

HTG; HTGS_PHASE0.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 61633)

AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

Homo sapiens chromosome 11, clone CTD-2380K20

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 61633)

* 10836 10935: gap of 100 bp
* 10936 11596: contig of 661 bp in length
* 11597 11696: gap of 100 bp
* 11697 12380: contig of 684 bp in length
* 12381 12480: gap of 100 bp
* 12481 13148: contig of 668 bp in length
* 13149 13248: gap of 100 bp
* 13249 13915: contig of 667 bp in length
* 13916 14015: gap of 100 bp
* 14016 14693: contig of 678 bp in length
* 14694 14793: gap of 100 bp
* 14794 15464: contig of 671 bp in length
* 15465 15564: gap of 100 bp
* 15565 16255: contig of 691 bp in length
* 16256 16355: gap of 100 bp
* 16356 17032: contig of 677 bp in length
* 17033 17132: gap of 100 bp
* 17133 17814: contig of 682 bp in length
* 17815 17914: gap of 100 bp
* 17915 18594: contig of 680 bp in length
* 18595 18694: gap of 100 bp
* 18695 19371: contig of 677 bp in length
* 19372 19471: gap of 100 bp
* 19472 20135: contig of 664 bp in length
* 20136 20235: gap of 100 bp
* 20236 20913: contig of 678 bp in length
* 20914 21013: gap of 100 bp
* 21014 21661: contig of 648 bp in length
* 21662 21761: gap of 100 bp
* 21762 22438: contig of 677 bp in length
* 22439 22538: gap of 100 bp
* 22539 23244: contig of 706 bp in length
* 23245 23344: gap of 100 bp
* 23345 24006: contig of 662 bp in length
* 24007 24106: gap of 100 bp
* 24107 24800: contig of 694 bp in length
* 24801 24900: gap of 100 bp
* 24901 25593: contig of 693 bp in length
* 25594 25693: gap of 100 bp
* 25694 26386: contig of 693 bp in length
* 26387 26486: gap of 100 bp
* 26487 27136: contig of 650 bp in length
* 27137 27236: gap of 100 bp
* 27237 27934: contig of 698 bp in length
* 27935 28034: gap of 100 bp
* 28035 28709: contig of 675 bp in length
* 28710 28809: gap of 100 bp
* 28810 29480: contig of 671 bp in length
* 29481 29580: gap of 100 bp
* 29581 30252: contig of 672 bp in length
* 30253 30352: gap of 100 bp
* 30353 31017: contig of 665 bp in length
* 31018 31117: gap of 100 bp
* 31118 31852: contig of 735 bp in length
* 31853 31952: gap of 100 bp
* 31953 32642: contig of 690 bp in length
* 32643 32742: gap of 100 bp
* 32743 33439: contig of 697 bp in length
* 33440 33539: gap of 100 bp
* 33540 34229: contig of 690 bp in length
* 34230 34329: gap of 100 bp
* 34330 35029: contig of 700 bp in length
* 35030 35129: gap of 100 bp
* 35130 35825: contig of 696 bp in length
* 35826 35925: gap of 100 bp
* 35926 36606: contig of 681 bp in length
* 36607 36706: gap of 100 bp
* 36707 37362: contig of 656 bp in length
* 37363 37462: gap of 100 bp
* 37463 38129: contig of 667 bp in length
* 38130 38229: gap of 100 bp
* 38230 38903: contig of 674 bp in length
* 38904 39003: gap of 100 bp

* 39004 39681: contig of 678 bp in length
* 39682 39781: gap of 100 bp
* 39782 40469: contig of 688 bp in length
* 40470 40569: gap of 100 bp
* 40570 41258: contig of 689 bp in length
* 41259 41358: gap of 100 bp
* 41359 42053: contig of 695 bp in length
* 42054 42153: gap of 100 bp
* 42154 42805: contig of 652 bp in length
* 42806 42905: gap of 100 bp
* 42906 43559: contig of 654 bp in length
* 43560 43659: gap of 100 bp
* 43660 44324: contig of 665 bp in length
* 44325 44424: gap of 100 bp
* 44425 45087: contig of 663 bp in length
* 45088 45187: gap of 100 bp
* 45188 45878: contig of 691 bp in length
* 45879 45978: gap of 100 bp
* 45979 46645: contig of 667 bp in length
* 46646 46745: gap of 100 bp
* 46746 47439: contig of 694 bp in length
* 47440 47539: gap of 100 bp
* 47540 48230: contig of 691 bp in length
* 48231 48330: gap of 100 bp
* 48331 49027: contig of 697 bp in length
* 49028 49127: gap of 100 bp
* 49128 49798: contig of 671 bp in length
* 49799 49898: gap of 100 bp
* 49899 50566: contig of 668 bp in length
* 50567 50666: gap of 100 bp
* 50667 51354: contig of 688 bp in length
* 51355 51454: gap of 100 bp
* 51455 52137: contig of 683 bp in length
* 52138 52237: gap of 100 bp
* 52238 52905: contig of 668 bp in length
* 52906 53005: gap of 100 bp
* 53006 53685: contig of 680 bp in length
* 53686 53785: gap of 100 bp
* 53786 54487: contig of 702 bp in length

Query Match

Best Local Similarity 84.8%; Score 17.8; DB 2; Length 61633;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgcgtctgtggggggg 21

Db 35018 GGGGAGGTCGTCTGGGGGGG 34998

RESULT 12

AC068923

LOCUS

AC068923 136120 bp DNA linear PLN 05-JAN-2002
Oryza sativa chromosome 10 BAC OSNBA0017E08 genomic sequence,
complete sequence.

AC068923

ACCESSION

VERSION

AC068923.11 GI:17298629

KEYWORDS

HTG

SOURCE

Oryza sativa.

ORGANISM

Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 136120)

Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Moffat,K.S., Hill,J.N.,

Gausberger,K., Brenner,M., Burgess,S., Hance,M., Shvartsbeyn,M.,

Tsitrin,T., Riggs,F., Hsiao,J., Zismann,V., Blunt,S., Pal,G.,

VanAken,S.E., Utterback,T.R., Feldblyum,T.V., Kalb,E.,

Quackenbush,J., Salzberg,S.L., White,O. and Fraser,C.M.

Oryza sativa chromosome 10 BAC OSNBA0017E08 genomic sequence

Unpublished

2 (bases 1 to 136120)

Buell,R.

Direct Submission

TITLE


```

QPGRAEADGFRYNKRVVRYIILFSKSEIPMAKEDQEIIMDLADLSTLPPEFTEDL
DYSSLYSGFRYPLOMPYRSLHKKRPIGDAKFQVPVVMGREDYVYKFGVESAI
KDGTHERHAPDMKITIPEGSHFVQEQFPDYVYNELLAFKDHDPVK"
complement(16539..16563)
/rpt_family="CCGCG"n"
complement(16966..17003)
/rpt_family="AT-rich"
complement(join(<18641..18896,18984..19219,19392..19458,
19981..20198,20292..20378))
/gene="OSJNBa0017E08.11"
complement(18641..20378)
/gene="OSJNBa0017E08.11"
/note="similar to epoxide hydrolase GB:AA02006 GI:1354849
(Nicotiana tabacum)"
complement(join(18641..18896,18984..19219,19392..19458,
19981..20198,20292..20378))
/gene="OSJNBa0017E08.11"
/codon_start=1
/product="putative epoxide hydrolase"
/protein_id="AAL58278.1"
/db_xref="GI:18071419"
/translation="NATAEQOOQIEHVLPVRGSPSTSRKPAKASGRWCSTGSRRY
GTRGATRCSPPPGTAPSRRTGAGTSGFTSRSSRRRTTSSRISPSPTSPCPR
PELLGRTLPQOLIVLLSVIQIPGRADFGKDIRRVVRYIILFSRNEIPIAKED
QEIMDLADLSTLPPEFTEDLDVYSSLYSGFRYPLOMPYRSLHKKRPIGDAKFQV
PVVVMGREDYVYKFGSHFVQEQFPDYVYNELLAFKDHDPVK"
LSFLKDHDPVAV"
complement(20554..20628)
/rpt_family="(TA)n"
complement(21642..21669)
/rpt_family="AT-rich"
complement(join(<21983..22238,22326..22561,22644..22805,

```

```

Query Match      84.8%; Score 17.8; DB 8; Length 136120;
Best Local Similarity 90.5%; Pred. No. 1e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggagcagcgtcgtggggggg 21
|| ||||| ||||| |||||
Db 115933 GGTGAGCTGCTGCTGGGGGG 115953

```

```

RESULT 13
AC095209
LOCUS      157246 bp      DNA      linear      HTG 20-DEC-2001
DEFINITION Rattus norvegicus clone CH230-9D23, ** SEQUENCING IN PROGRESS **
AC095209
VERSION    94 unordered pieces.
KEYWORDS   HTG; HTGS; PHASE1.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 157246)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Binage,K., Blankenburg,K., Bonin,D., Bouck,J.,
Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Covle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,

```

```

Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Kratovick,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Katsic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,M., Nickerson,E., Nwokenkwo,S.,
Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojoubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.R.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 157246)
Worley,K.C.
Submitted (16-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15625763.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDHV
Center clone name: CH230-9D23
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 117566 bases at least Q40
Consensus quality: 126981 bases at least Q30
Consensus quality: 133584 bases at least Q20
Estimated insert size: 84204; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.1x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 94 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 3299: contig of 3298 bp in length
* 3398: gap of unknown length
* 3399: contig of 2437 bp in length
* 5836: gap of unknown length
* 5935: gap of unknown length
* 9609: contig of 3674 bp in length
* 9709: gap of unknown length
* 9710: contig of 2307 bp in length
* 12016: contig of 2193 bp in length
* 12117: gap of unknown length
* 12117: contig of 2193 bp in length
* 14310: gap of unknown length
* 14409: contig of 2254 bp in length
* 16663: gap of unknown length
* 16664: contig of 1561 bp in length
* 18325: gap of unknown length

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

18425 20791: contig of 2367 bp in length
 20792 20891: gap of unknown length
 20892 22208: contig of 1317 bp in length
 22209 22308: gap of unknown length
 22309 23677: contig of 1369 bp in length
 23678 23777: gap of unknown length
 23778 26517: contig of 2740 bp in length
 26518 26617: gap of unknown length
 26619 28924: contig of 2307 bp in length
 28925 29024: gap of unknown length
 29026 30860: contig of 1836 bp in length
 30861 30960: gap of unknown length
 30961 32855: contig of 1895 bp in length
 32856 32955: gap of unknown length
 32956 35481: contig of 2526 bp in length
 35482 35581: gap of unknown length
 35582 37274: contig of 1693 bp in length
 37275 37374: gap of unknown length
 37375 38663: contig of 1289 bp in length
 38664 38763: gap of unknown length
 38764 40690: contig of 1927 bp in length
 40691 40790: gap of unknown length
 40791 42883: contig of 2093 bp in length
 42884 44557: gap of unknown length
 44558 44657: gap of unknown length
 44659 46940: contig of 2283 bp in length
 46941 47040: gap of unknown length
 47041 48478: contig of 1438 bp in length
 48479 48578: gap of unknown length
 48579 50529: contig of 1951 bp in length
 50530 50629: gap of unknown length
 50630 52404: contig of 1775 bp in length
 52405 52504: gap of unknown length
 52505 53708: contig of 1204 bp in length
 53709 53808: gap of unknown length
 53809 55212: contig of 1404 bp in length
 55213 55312: gap of unknown length
 55313 56526: contig of 1214 bp in length
 56527 56626: gap of unknown length
 56627 58151: contig of 1525 bp in length
 58152 58251: gap of unknown length
 58252 60034: contig of 1783 bp in length
 60035 60134: gap of unknown length
 60135 61711: contig of 1577 bp in length
 61712 61811: gap of unknown length
 61812 63314: contig of 1503 bp in length
 63315 63414: gap of unknown length
 63415 64738: contig of 1324 bp in length
 64739 64838: gap of unknown length
 64839 65971: contig of 1133 bp in length
 65972 66071: gap of unknown length
 66072 68054: contig of 1983 bp in length
 68055 68154: gap of unknown length
 68155 69615: contig of 1461 bp in length
 69616 69715: gap of unknown length
 69716 71212: contig of 1497 bp in length
 71213 73385: contig of 2073 bp in length
 73386 73486: gap of unknown length
 73487 75145: contig of 1660 bp in length
 75146 76847: contig of 1602 bp in length
 76848 76947: gap of unknown length
 76948 78558: contig of 1611 bp in length
 78559 78658: gap of unknown length
 78659 80361: contig of 1703 bp in length
 80362 80461: gap of unknown length
 80462 82509: contig of 2048 bp in length
 82510 82610: gap of unknown length
 82610 84259: contig of 1650 bp in length
 84260 84359: gap of unknown length
 84360 85826: contig of 1467 bp in length

85827 85926: gap of unknown length
 85927 87333: contig of 1407 bp in length
 87334 87434: gap of unknown length
 87435 88649: contig of 1216 bp in length
 88650 88749: gap of unknown length
 88750 90431: contig of 1682 bp in length
 90432 90531: gap of unknown length
 90532 91847: contig of 1316 bp in length
 91848 91947: gap of unknown length
 91949 93463: contig of 1516 bp in length
 93464 93563: gap of unknown length
 93564 94927: contig of 1364 bp in length
 94928 95027: gap of unknown length
 95028 96552: contig of 1525 bp in length
 96553 96652: gap of unknown length
 96653 98176: contig of 1524 bp in length
 98177 98276: gap of unknown length
 98277 99608: contig of 1332 bp in length
 99609 99708: gap of unknown length
 99709 101314: contig of 1606 bp in length
 101315 101414: gap of unknown length
 101415 102640: contig of 1226 bp in length
 102641 102740: gap of unknown length
 102741 104243: contig of 1503 bp in length
 104244 104343: gap of unknown length

Query Match

Best Local Similarity 84.8%; Score 17.8; DB 2; Length 157246;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgcctgctggggggg 21

DB 112300 GGGGACGGCTGTGGGGGG 112320

RESULT 14

AC105838/c

LOCUS

DEFINITION

AC105838

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC105838 160867 bp DNA linear HTG 10-JAN-2002
 Rattus norvegicus clone CH230-46C20, *** SEQUENCING IN PROGRESS
 ***, 70 unordered pieces.

AC105838

AC105838.1 GI:18104745

HTG: HTGS_PHASE1.

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 160867)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
 Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
 Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Denn,A.L., Ding,Y., Dinh,H., Douthwaite,K.J., Draper,H.,
 Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,
 Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,K., Jolivet,S.,
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
 Lounsged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,

```

Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,S., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogih,M., Okwono,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaika,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Washington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 160867)
Worley,K.C.
Direct Submission
Submitted (10-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNVX
Center clone name: CH230-46C20
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 125582 bases at least Q40
Consensus quality: 134475 bases at least Q30
Consensus quality: 139444 bases at least Q20
Estimated insert size: 121068; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.1x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 70 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
*
* 1 6587: contig of 6587 bp in length
* 6588 6687: gap of unknown length
* 6688 13446: contig of 6659 bp in length
* 13447 13446: gap of unknown length
* 13447 18658: contig of 5212 bp in length
* 18659 18758: gap of unknown length
* 18759 23839: contig of 5081 bp in length
* 23840 23939: gap of unknown length
* 23940 23969: contig of 5430 bp in length
* 23970 29469: gap of unknown length
* 29470 33207: contig of 3738 bp in length
* 33208 33307: gap of unknown length
* 33308 37028: contig of 3721 bp in length
* 37029 37128: gap of unknown length
* 37129 39332: contig of 2204 bp in length
* 39333 39432: gap of unknown length
* 39433 43860: contig of 4428 bp in length
* 43861 43960: gap of unknown length
* 43961 48285: contig of 4325 bp in length
* 48286 48385: gap of unknown length
* 48386 50711: contig of 2326 bp in length
* 50712 50811: gap of unknown length

```

```

50812 53428: contig of 2617 bp in length
53429 53528: gap of unknown length
53529 5426: contig of 1397 bp in length
5426 55025: gap of unknown length
55025 58141: contig of 3116 bp in length
58141 58241: gap of unknown length
58241 61585: contig of 3344 bp in length
61585 61685: gap of unknown length
61685 63685: contig of 2000 bp in length
63685 63785: gap of unknown length
63785 66332: contig of 2447 bp in length
66332 66333: gap of unknown length
66333 68961: contig of 2629 bp in length
68961 71380: contig of unknown length
71380 71480: contig of 2319 bp in length
71480 75439: contig of unknown length
75439 75440: contig of 3959 bp in length
75440 77499: contig of unknown length
77499 77599: gap of unknown length
77599 79007: contig of 1408 bp in length
79007 79107: gap of unknown length
80462 80462: contig of 1355 bp in length
80462 82640: contig of 2078 bp in length
82640 82641: gap of unknown length
82641 85396: contig of 2656 bp in length
85396 85496: gap of unknown length
85496 87789: contig of 2293 bp in length
87789 88959: contig of unknown length
88959 91799: gap of unknown length
91799 91899: contig of 2740 bp in length
91899 93876: contig of 1977 bp in length
93876 93976: gap of unknown length
93976 95422: contig of 1346 bp in length
95422 97309: contig of unknown length
97309 97409: gap of unknown length
97409 100533: contig of 2644 bp in length
100533 100153: gap of unknown length
100153 102572: contig of 2419 bp in length
102572 102673: gap of unknown length
102673 105229: contig of 2557 bp in length
105229 105329: gap of unknown length
105329 107336: contig of 2207 bp in length
107336 107637: gap of unknown length
107637 109788: contig of 2152 bp in length
109788 111257: contig of 1369 bp in length
111257 111357: gap of unknown length
111357 112859: contig of 1502 bp in length
112859 112960: gap of unknown length
112960 114728: contig of 1769 bp in length
114728 114828: gap of unknown length
114828 116778: contig of 1950 bp in length
116778 116878: gap of unknown length
116878 118806: contig of 1928 bp in length
118806 118906: gap of unknown length
118906 120388: contig of 1482 bp in length
120388 120488: gap of unknown length
120488 121578: contig of 1090 bp in length
121578 121678: gap of unknown length
121678 123382: contig of 1704 bp in length
123382 123482: gap of unknown length
123482 124551: contig of 1069 bp in length
124551 124651: gap of unknown length
124651 126461: contig of 1810 bp in length
126461 126561: gap of unknown length
126561 127898: contig of 1337 bp in length
127898 127955: gap of unknown length
127955 129755: contig of 1757 bp in length

```

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

* 129756 129855: gap of unknown length
 * 129856 131417: contig of 1562 bp in length
 * 131418 131517: gap of unknown length
 * 131518 132520: contig of 1003 bp in length
 * 132521 132620: gap of unknown length
 * 132621 133913: contig of 1293 bp in length
 * 133914 134013: gap of unknown length
 * 134014 135386: contig of 1373 bp in length
 * 135387 135486: gap of unknown length
 * 135487 136882: contig of 1396 bp in length
 * 136883 136982: gap of unknown length
 * 136983 138334: contig of 1352 bp in length
 * 138335 138434: gap of unknown length
 * 138435 139804: contig of 1370 bp in length
 * 139805 139904: gap of unknown length
 * 139905 141196: contig of 1292 bp in length
 * 141197 141296: gap of unknown length
 * 141297 142755: contig of 1459 bp in length

Query Match 84.8%; Score 17.8; DB 2; Length 160867;
 Best Local Similarity 90.5%; Pred. No. 9.8e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacacgcctgctg999999 21

Db 37003 GGGGACCAAGCTCTGGGGGG 36983

RESULT 15

AC094985

LOCUS

DEFINITION

Rattus norvegicus clone CH230-6E13, *** SEQUENCING IN PROGRESS ***

84 unordered pieces.

AC094985 2 GI:17941787

HTG; HTGS_PHASE1.

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 166036)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Banks,T., Barbara,J.,
 Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
 Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Denn,A.L., Ding,Y., Dinh,H., Douthwaite,K.J., Draper,H.,
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,
 Hollins,B., Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
 Loulsegue,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
 Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
 Morgan,N., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
 Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
 Quiles,M., Remy,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
 Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,

Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
 Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
 Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
 Worley,K., Wu,C., Wu,Y., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 166036)

Worley,K.C.

Direct Submission

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:15624822.

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GBWV

Center clone name: CH230-6E13

----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to

findPhrapList

Consensus quality: 112466 bases at least Q40

Consensus quality: 123656 bases at least Q30

Consensus quality: 130581 bases at least Q20

Estimated insert size: 97933; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 1.2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 84 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 3626: contig of 3626 bp in length

* 3627 3726: gap of unknown length

* 3727 7568: contig of 3842 bp in length

* 7569 7668: gap of unknown length

* 7669 12337: contig of 4669 bp in length

* 12338 12437: gap of unknown length

* 12438 15712: contig of 3275 bp in length

* 15713 15812: gap of unknown length

* 15813 19342: contig of 3530 bp in length

* 19343 19442: gap of unknown length

* 19443 22577: contig of 3135 bp in length

* 22578 22677: gap of unknown length

* 22678 24726: contig of 2049 bp in length

* 24727 24826: gap of unknown length

* 24827 29286: contig of 4460 bp in length

* 29287 29387: gap of unknown length

* 29387 33090: contig of 3704 bp in length

* 33091 33190: gap of unknown length

* 33191 35988: contig of 2798 bp in length

* 35989 36088: gap of unknown length

* 36089 38247: contig of 2159 bp in length

* 38248 38348: gap of unknown length

* 38348 41204: contig of 2857 bp in length

* 41205 41304: gap of unknown length

* 41305 43271: contig of 1967 bp in length

* 43272 43371: gap of unknown length

* 43372 46578: contig of 3207 bp in length

* 46578 46678: gap of unknown length

* 46579

* 46679 48809: contig of 2131 bp in length
* 48810 48909: gap of unknown length
* 48910 51058: contig of 2149 bp in length
* 51059 51158: gap of unknown length
* 51159 53838: contig of 2680 bp in length
* 53839 53938: gap of unknown length
* 53939 56061: contig of 2123 bp in length
* 56062 56161: gap of unknown length
* 56162 57868: contig of 1707 bp in length
* 57869 57969: gap of unknown length
* 57969 60202: contig of 2234 bp in length
* 60203 60302: gap of unknown length
* 60303 62705: contig of 2403 bp in length
* 62706 62805: gap of unknown length
* 62806 64569: contig of 1764 bp in length
* 64570 64669: gap of unknown length
* 64670 67010: contig of 2341 bp in length
* 67011 67110: gap of unknown length
* 67111 69044: contig of 1934 bp in length
* 69045 69144: gap of unknown length
* 69145 70499: contig of 1355 bp in length
* 70500 70599: gap of unknown length
* 70600 72974: contig of 2375 bp in length
* 72975 73074: gap of unknown length
* 73075 75261: contig of 2187 bp in length
* 75262 75361: gap of unknown length
* 75362 76963: contig of 1602 bp in length
* 76964 77063: gap of unknown length
* 77064 79168: contig of 2105 bp in length
* 79169 79268: gap of unknown length
* 79269 80896: contig of 1628 bp in length
* 80897 83484: gap of unknown length
* 83485 86135: contig of 2551 bp in length
* 86136 86235: gap of unknown length
* 86236 87514: contig of 1279 bp in length
* 87515 87614: gap of unknown length
* 87615 89297: contig of 1683 bp in length
* 89298 89397: gap of unknown length
* 89398 91614: contig of 2217 bp in length
* 91615 91714: gap of unknown length
* 91715 93391: contig of 1677 bp in length
* 93392 93491: gap of unknown length
* 93492 95101: contig of 1610 bp in length
* 95102 95201: gap of unknown length
* 95202 96392: contig of 1191 bp in length
* 96393 96492: gap of unknown length
* 96493 99073: contig of 2581 bp in length
* 99074 99173: gap of unknown length
* 99174 100872: contig of 1699 bp in length
* 100873 100972: gap of unknown length
* 100973 102337: contig of 2265 bp in length
* 102338 103337: gap of unknown length
* 103338 105398: contig of 2061 bp in length
* 105399 105498: gap of unknown length
* 105499 106516: contig of 1018 bp in length
* 106517 106616: gap of unknown length
* 106617 108034: contig of 1418 bp in length
* 108035 108134: gap of unknown length
* 108135 109492: contig of 1358 bp in length
* 109493 109592: gap of unknown length
* 109593 110875: contig of 1283 bp in length
* 110876 110975: gap of unknown length
* 110976 112982: contig of 2007 bp in length
* 112983 113082: gap of unknown length
* 113083 114947: contig of 1865 bp in length
* 114948 115047: gap of unknown length
* 115048 116180: contig of 1133 bp in length
* 116181 116280: gap of unknown length
* 116281 117410: contig of 1130 bp in length
* 117411 117510: gap of unknown length
* 117511 119709: contig of 2199 bp in length

* 119710 119809: gap of unknown length
* 119810 121124: contig of 1315 bp in length
* 121125 121224: gap of unknown length
* 121225 122832: contig of 1608 bp in length
* 122833 122932: gap of unknown length
* 122933 124260: contig of 1328 bp in length
* 124261 124360: gap of unknown length
* 124361 125585: contig of 1225 bp in length
* 125586 125685: gap of unknown length
* 125686 127037: contig of 1352 bp in length
* 127038 127137: gap of unknown length

Query Match 84.8%; Score 17.8; DB 2; Length 166036;
Best Local Similarity 90.5%; Pred. No. 9.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ggggacgacgtcgtggggggg 21
||||| ||||||| |||||||
Db 162656 GGGGCCGACGTGCGGGGGGG 162676

Search completed: August 10, 2002, 02:59:11
Job time: 15757 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:24:27 ; Search time 1145.36 Seconds
(without alignments)
31.479 Million cell updates/sec

Title: US-09-672-126-37

Perfect score: 21

Sequence: 1 ggggacgacgtcggtggggg 21

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	ID	Description
1	21	100.0	21	22	AAF98767	Human IFN-alpha im
2	21	100.0	21	22	AAF98767	Immunostimulatory
3	19.4	92.4	22	22	AAF98740	Human IFN-alpha im
4	19.4	92.4	22	22	AAF99784	Immunostimulatory
5	18.4	87.6	20	22	AAF98880	Immunostimulatory
6	18.4	87.6	20	22	AAF99870	Immunostimulatory
7	17.8	84.8	22	22	AAF98751	Human IFN-alpha im
8	17.8	84.8	22	22	AAF99834	Immunostimulatory
9	16.8	80.0	20	22	AAF98735	Human IFN-alpha im

10	16.8	80.0	20	22	AAF98855	Poly-G immunostimu
11	16.8	80.0	20	22	AAF98871	Immunostimulatory
12	16.8	80.0	20	22	AAF98879	Immunostimulatory
13	16.8	80.0	20	22	AAF99704	Immunostimulatory
14	16.8	80.0	20	22	AAF99767	Immunostimulatory
15	16.8	80.0	20	22	AAF99868	Immunostimulatory
16	16.2	77.1	22	22	AAF98739	Human IFN-alpha im
17	16.2	77.1	22	22	AAF98741	Human IFN-alpha im
18	16.2	77.1	22	22	AAF99783	Immunostimulatory
19	16.2	77.1	22	22	AAF99785	Immunostimulatory
20	16.2	77.1	30	13	AAQ20870	Immunostimulatory
21	16.2	77.1	30	13	AAQ20873	Immunostimulatory
22	16.2	77.1	40	21	AAZ96150	Poly-nucleotide seq
23	16.2	77.1	149	21	AAZ21687	Human secreted pro
24	16.2	77.1	228	23	ABL25165	Drosophila melanog
25	16.2	77.1	462	22	AHL13268	Human cDNA clone (
26	16.2	77.1	748	21	AAFI5106	Trichoderma reesei
27	16.2	77.1	876	22	AAD17507	Sequence encoding
28	16.2	77.1	1053	8	AAAT70994	Human taste recept
29	16.2	77.1	1053	15	AAQ73489	DNA encoding gp63
30	16.2	77.1	1053	22	AAD10195	Pseudorabies virus
31	16.2	77.1	1053	22	AAS09831	Pseudorabies virus
32	16.2	77.1	1053	22	AAQ90826	PRV glycoprotein g
33	16.2	77.1	1386	18	AAAT50944	Mature endoglycoce
34	16.2	77.1	1473	18	AAAT50943	Full length endogl
35	16.2	77.1	1660	21	AAZ34782	Wheat sucrose phos
36	16.2	77.1	1974	22	ABA09482	Human secreted pro
37	16.2	77.1	2082	21	AAQ69802	Human breast tumou
38	16.2	77.1	2215	20	AAZ23924	HSV2 LAT DNA. Her
39	16.2	77.1	2288	23	ABL25164	Drosophila melanog
40	16.2	77.1	2553	22	AAD17509	Human taste recept
41	16.2	77.1	2559	24	AAQ97395	Human SAC1 gene cD
42	16.2	77.1	2739	24	ABK16615	Human G-coupled re
43	16.2	77.1	3384	22	AAS59854	Human novel cytoki
44	16.2	77.1	3384	23	AAQ77630	DNA encoding novel
45	16.2	77.1	3384	23	AAS88731	DNA encoding novel

ALIGNMENTS

RESULT 1

AAF98767

ID AAF98767 standard; DNA; 21 BP.

AC AAF98767;

DT 11-JUN-2001 (first entry)

DE Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 37.

DE Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;

KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.

XX Synthetic.

OS Key

FH modified_base

FT Location/Qualifiers

FT 1..22 a

FT /mod_base= "OTHER"

FT /note= "phosphorothioate linkage"

FT 16..20

FT modified_base

FT /tag= b

FT /mod_base= "OTHER"

FT /note= "phosphorothioate linkage"

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX Hartmann G, Bratzler RL, Krieg A;
 XX WPI; 2001-290487/30.
 DR Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 XX Claim 201; Page 103; 168pp; English.
 XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX Sequence 21 BP; 2 A; 3 C; 14 G; 2 T; 0 other;
 SQ

Query Match 100.0%; Score 21; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgacgtcgtggggggg 21
 Db 1 ggggacgacgtcgtggggggg 21
 ||||||||||||||||

RESULT 2
 AAF99873
 ID AAF99873 standard; DNA; 21 BP.
 XX AAF99873;
 XX 12-JUN-2001 (first entry)
 XX Immunostimulatory nucleic acid #989.
 DE Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 XX Immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX Synthetic.
 OS
 XX WO200122972-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 25-SEP-2000; 2000WO-US26383.
 PF
 XX 25-SEP-1999; 99US-0156113.
 PR
 XX 27-SEP-1999; 99US-0156135.
 PR
 XX 23-AUG-2000; 2000US-0227436.
 PR
 XX (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 PA
 XX Krieg AM, Schetter C, Vollmer J;
 XX WPI; 2001-273485/28.
 DR
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 PT
 XX Claim 101; Page 59; 338pp; English.

XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 XX Sequence 21 BP; 2 A; 3 C; 14 G; 2 T; 0 other;
 SQ

Query Match 100.0%; Score 21; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggggacgacgtcgtggggggg 21
 Db 1 ggggacgacgtcgtggggggg 21
 ||||||||||||||||

RESULT 3
 AAF98740
 ID AAF98740 standard; DNA; 22 BP.
 XX AAF98740;
 XX 11-JUN-2001 (first entry)
 DT Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 10.
 DE Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
 XX viral infection; phosphorothioate backbone; palindrome; cancer; ds.
 KW
 KW Synthetic.
 OS
 XX Key Location/Qualifiers
 FH modified_base 1..2
 FT /*tag= a
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 FT modified_base 17..21
 FT /*tag= b
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 XX WO200122990-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 27-SEP-2000; 2000WO-US26527.
 PF
 XX 27-SEP-1999; 99US-0156147.
 PR
 XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 PA
 XX Hartmann G, Bratzler RL, Krieg A;
 PI
 XX WPI; 2001-290487/30.
 DR
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 PT
 XX Claim 201; Page 103; 168pp; English.

XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX Sequence 22 BP; 2 A; 4 C; 14 G; 2 T; 0 other;
 SQ

Query Match 92.4%; Score 19.4; DB 22; Length 22;
 Best Local Similarity 95.2%; Pred. No. 34;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgacgtcgtcgggggg 21
 |||||
 Db 2 ggggacgacgtcgtcgggggg 22

RESULT 4
 AAF99784
 ID AAF99784 standard; DNA; 22 BP.
 AC AAF99784;
 XX
 XX 12-JUN-2001 (first entry)
 DT
 XX Immunostimulatory nucleic acid #900.
 DE Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX Synthetic.
 OS
 XX WO200122972-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 25-SEP-2000; 2000WO-US26383.
 PF
 XX 25-SEP-1999; 99US-0156113.
 PR 27-SEP-1999; 99US-0156135.
 PR 23-AUG-2000; 2000US-0227436.
 XX
 XX (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 XX
 XX Krieg AM, Schetter C, Vollmer J;
 PI
 XX WPI; 2001-273485/28.
 DR
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids -
 PT
 XX Claim 101; Page 57; 338pp; English.
 PS
 XX The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g., herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g., toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a

CC Th2 to a Th1 immune response and to activate immune cells.
 CC Note: the present sequence may have a phosphorothioate backbone.
 XX
 SQ Sequence 22 BP; 2 A; 4 C; 14 G; 2 T; 0 other;
 SQ

Query Match 92.4%; Score 19.4; DB 22; Length 22;
 Best Local Similarity 95.2%; Pred. No. 34;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ggggacgacgtcgtcgggggg 21
 |||||
 Db 2 ggggacgacgtcgtcgggggg 22

RESULT 5
 AAF98880
 ID AAF98880 standard; DNA; 20 BP.
 AC AAF98880;
 XX
 XX 11-JUN-2001 (first entry)
 DT
 XX Immunostimulatory nucleic acid assay control oligo SEQ ID NO: 161.
 DE Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
 KW viral infection; phosphorothioate backbone; palindromes; cancer; ds.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH modified_base 1..2
 FT /tag= a
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 FT modified_base 15..19
 FT /tag= b
 FT /mod_base= "OTHER"
 FT /note= "phosphorothioate linkage"
 XX
 PN WO200122990-A2.
 XX
 XX 05-APR-2001.
 PD
 XX 27-SEP-2000; 2000WO-US26527.
 PF
 XX 27-SEP-1999; 99US-0156147.
 PR
 XX (COLE-) COLEY PHARM GROUP INC.
 PA (IOWA) UNIV IOWA RES FOUND.
 XX
 XX Hartmann G, Bratzler RL, Krieg A;
 PI
 XX WPI; 2001-290487/30.
 DR
 XX Improving the efficacy of treatments involving the administration of
 PT interferon-alpha by co-administering an isolated immunostimulatory
 PT nucleic acid -
 PT
 XX Example 17; Page 167; 168pp; English.
 PS
 XX The present invention describes an improvement to a method requiring the
 CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.
 XX
 XX Sequence 20 BP; 1 A; 3 C; 13 G; 3 T; 0 other;
 SQ


```
AC AAF99834;
XX 12-JUN-2001 (first entry)
XX Immunostimulatory nucleic acid #950.
DE
XX Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS Synthetic.
XX
XX WO200122972-A2.
XX
XX 05-APR-2001.
XX
XX 25-SEP-2000; 2000WO-US26383.
XX
XX 25-SEP-1999; 99US-0156113.
XX
XX 27-SEP-1999; 99US-0156135.
XX
XX 23-AUG-2000; 2000US-0227436.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX
XX (COLE-) COLEY PHARM GMBH.
XX
XX Krieg AM, Schetter C, Vollmer J;
XX
XX WPI; 2001-273485/28.
XX
XX Vaccinating against tumors, infectious diseases, allergies and asthma;
XX using immunostimulatory Py-rich and TG nucleic acids -
XX
XX Claim 101; Page 58; 338pp; English.
XX
XX The present invention relates to a method for stimulating an immune
XX response. The method comprises administering an immunostimulatory nucleic
XX acid to a non-rodent subject in sufficient quantity to stimulate an
XX immune response. The present sequence is one such immunostimulatory
XX nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
XX (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
XX against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
XX and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
XX haemophilus, campylobacter, clostridium, Escherichia coli and/or
XX staphylococcus), fungal antigens and/or parasitic antigens. The method is
XX also useful for preventing cancer, asthma, infectious disease, allergy or
XX immune deficiency. The present sequence can also be used to redirect a
XX Th2 to a Th1 immune response and to activate immune cells.
XX Note: the present sequence may have a phosphorothioate backbone.
XX
XX Sequence 22 BP; 3 A; 3 C; 13 G; 3 T; 0 other;
SQ
```

```
Query Match 84.8%; Score 17.8; DB 22; Length 22;
Best Local Similarity 90.5%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ggggacgacgtcgtggggggg 21
   ||| ||||| ||||| |||||
DB 2 ggggacgacgtcgtggggggg 22
```

```
RESULT 9
AAF98735
ID AAF98735 standard; DNA; 20 BP.
XX
XX AAF98735;
XX
XX 11-JUN-2001 (first entry)
XX
XX Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 5.
XX
XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
```

```
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FT modified_base 1..20
FT /*tag= a
FT /*mod_base= "OTHER"
FT /*note= "phosphorothioate linkage"
XX
XX WO200122990-A2.
XX
XX 05-APR-2001.
XX
XX 27-SEP-2000; 2000WO-US26527.
XX
XX 27-SEP-1999; 99US-0156147.
XX
XX (COLE-) COLEY PHARM GROUP INC.
XX (IOWA ) UNIV IOWA RES FOUND.
XX
XX Hartmann G, Bratzler RL, Krieg A;
XX
XX WPI; 2001-290487/30.
XX
XX Improving the efficacy of treatments involving the administration of
XX interferon-alpha by co-administering an isolated immunostimulatory
XX nucleic acid -
XX
XX Claim 201; Page 103; 168pp; English.
XX
XX The present invention describes an improvement to a method requiring the
XX administration of interferon alpha (IFN-alpha), involving administering
XX an immunostimulatory nucleic acid (ISNA). The sequences of a number of
XX such nucleic acids are also provided. These may comprise oligonucleotides
XX with phosphorothioate backbones, palindromes, or G-rich sequences. The
XX sequences of the invention are useful in the treatment of proliferative
XX diseases, such as cancers, and viral infections. The present sequence is
XX an example of an immunostimulatory oligonucleotide.
XX
XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;
SQ
```

```
Query Match 80.0%; Score 16.8; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 3.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ggggacgacgtcgtggggggg 20
   ||| ||||| ||||| |||||
DB 1 ggggacgacgtcgtggggggg 20
```

```
RESULT 10
AAF98855
ID AAF98855 standard; DNA; 20 BP.
XX
XX AAF98855;
XX
XX 11-JUN-2001 (first entry)
XX
XX Poly-G immunostimulatory nucleic acid SEQ ID NO: 136.
XX
XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
XX viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
XX Synthetic.
XX
XX WO200122990-A2.
XX
XX 05-APR-2001.
XX
XX 27-SEP-2000; 2000WO-US26527.
XX
```

```
PR 27-SEP-1999; 99US-0156147.
XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA ) UNIV IOWA RES FOUND.
XX
PI Hartmann G, Bratzler RL, Krieg A;
XX WPI; 2001-290487/30.
XX
XX Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
XX Disclosure; Page 24; 168pp; English.
XX
XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX
XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;
SQ
Query Match 80.0%; Score 16.8; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 3.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ggggacgcgtcggtggggg 20
DB 1 ggggacgcgtcggtggggg 20
RESULT 11
AAF98871
ID AAF98871 standard; DNA; 20 BP.
XX
XX AAF98871;
AC
XX
XX 11-JUN-2001 (first entry)
DT
XX
XX Immunostimulatory nucleic acid assay control oligo SEQ ID NO: 152.
DE
XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
XX Synthetic.
OS
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag= a
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
FT modified_base 15..19
FT /tag= b
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
XX
XX WO200122990-A2.
PN
XX
XX 05-APR-2001.
PD
XX
XX 27-SEP-2000; 2000WO-US26527.
PF
XX
XX 27-SEP-1999; 99US-0156147.
PR
XX
XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA ) UNIV IOWA RES FOUND.
XX
XX Hartmann G, Bratzler RL, Krieg A;
PI
XX WPI; 2001-290487/30.
DR
XX
XX Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
XX Example 17; Page 166; 168pp; English.
PS
XX The present invention describes an improvement to a method requiring the
```

```
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
XX Example 17; Page 163; 168pp; English.
XX
XX The present invention describes an improvement to a method requiring the
CC administration of interferon alpha (IFN-alpha), involving administering
CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC such nucleic acids are also provided. These may comprise oligonucleotides
CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC sequences of the invention are useful in the treatment of proliferative
CC diseases, such as cancers, and viral infections. The present sequence is
CC an example of an immunostimulatory oligonucleotide.
XX
XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;
SQ
Query Match 80.0%; Score 16.8; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 3.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ggggacgcgtcggtggggg 20
DB 1 ggggacgcgtcggtggggg 20
RESULT 12
AAF98879
ID AAF98879 standard; DNA; 20 BP.
XX
XX AAF98879;
AC
XX
XX 11-JUN-2001 (first entry)
DT
XX
XX Immunostimulatory nucleic acid assay control oligo SEQ ID NO: 160.
DE
XX Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
XX Synthetic.
OS
FH Key Location/Qualifiers
FT modified_base 1..2
FT /tag= a
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
FT modified_base 15..19
FT /tag= b
FT /mod_base= "OTHER"
FT /note= "phosphorothioate linkage"
XX
XX WO200122990-A2.
PN
XX
XX 05-APR-2001.
PD
XX
XX 27-SEP-2000; 2000WO-US26527.
PF
XX
XX 27-SEP-1999; 99US-0156147.
PR
XX
XX (COLE-) COLEY PHARM GROUP INC.
PA (IOWA ) UNIV IOWA RES FOUND.
XX
XX Hartmann G, Bratzler RL, Krieg A;
PI
XX WPI; 2001-290487/30.
DR
XX
XX Improving the efficacy of treatments involving the administration of
PT interferon-alpha by co-administering an isolated immunostimulatory
PT nucleic acid -
XX
XX Example 17; Page 166; 168pp; English.
PS
XX The present invention describes an improvement to a method requiring the
```

CC administration of interferon alpha (IFN-alpha), involving administering
 CC an immunostimulatory nucleic acid (ISNA). The sequences of a number of
 CC such nucleic acids are also provided. These may comprise oligonucleotides
 CC with phosphorothioate backbones, palindromes, or G-rich sequences. The
 CC sequences of the invention are useful in the treatment of proliferative
 CC diseases, such as cancers, and viral infections. The present sequence is
 CC an example of an immunostimulatory oligonucleotide.

XX Sequence 20 BP; 0 A; 3 C; 13 G; 4 T; 0 other;

Query Match 80.0%; Score 16.8; DB 22; Length 20;

Best Local Similarity 90.0%; Pred. No. 3.9e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gggagcagctgctggggggg 21

||||| ||||| ||||| |||||

Db 1 gggctgctgctggggggg 20

RESULT 13

AAF99704

ID AAF99704 standard; DNA; 20 BP.

XX AC AAF99704;

DT 12-JUN-2001 (first entry)

DE Immunostimulatory nucleic acid #820.

Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 immunostimulatory; tumour; viral infection; bacterial infection;
 fungal infection; parasitic infection; cancer; asthma;
 infectious disease; allergy; immune deficiency; phosphorothioate; ss.

OS Synthetic.

XX WO200122972-A2.

XX 05-APR-2001.

XX 25-SEP-2000; 2000WO-US26383.

XX 25-SEP-1999; 99US-0156113.

XX 27-SEP-1999; 99US-0156135.

XX 23-AUG-2000; 2000US-0227436.

XX (IOWA) UNIV IOWA RES FOUND.

PA (COLE-) COLEY PHARM GMBH.

XX Krieg AM, Schetter C, Vollmer J;

XX WPI; 2001-273485/28.

XX Vaccinating against tumors, infectious diseases, allergies and asthma

using immunostimulatory Py-rich and TG nucleic acids -

Claim 101; Page 56; 338pp; English.

The present invention relates to a method for stimulating an immune

response. The method comprises administering an immunostimulatory nucleic

acid to a non-rodent subject in sufficient quantity to stimulate an

immune response. The present sequence is one such immunostimulatory

nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich

(py-rich) or thymidine (T) rich. The method is used to vaccinate subjects

against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae

and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,

haemophilus, campylobacter, clostridium, Escherichia coli and/or

staphylococcus), fungal antigens and/or parasitic antigens. The method is

also useful for preventing cancer, asthma, infectious disease, allergy or

immune deficiency. The present sequence can also be used to redirect a

Th2 to a Th1 immune response and to activate immune cells.

Note: the present sequence may have a phosphorothioate backbone.

XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 80.0%; Score 16.8; DB 22; Length 20;

Best Local Similarity 90.0%; Pred. No. 3.9e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggagcagctgctggggggg 20

||||| ||||| ||||| |||||

Db 1 ggggtcgacgtcgagggggg 20

RESULT 14

AAF99767

ID AAF99767 standard; DNA; 20 BP.

XX AC AAF99767;

DT 12-JUN-2001 (first entry)

DE Immunostimulatory nucleic acid #883.

Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 immunostimulatory; tumour; viral infection; bacterial infection;
 fungal infection; parasitic infection; cancer; asthma;
 infectious disease; allergy; immune deficiency; phosphorothioate; ss.

OS Synthetic.

XX WO200122972-A2.

XX 05-APR-2001.

XX 25-SEP-2000; 2000WO-US26383.

XX 25-SEP-1999; 99US-0156113.

XX 27-SEP-1999; 99US-0156135.

XX 23-AUG-2000; 2000US-0227436.

XX (IOWA) UNIV IOWA RES FOUND.

PA (COLE-) COLEY PHARM GMBH.

XX Krieg AM, Schetter C, Vollmer J;

XX WPI; 2001-273485/28.

XX Vaccinating against tumors, infectious diseases, allergies and asthma

using immunostimulatory Py-rich and TG nucleic acids -

Claim 101; Page 57; 338pp; English.

The present invention relates to a method for stimulating an immune

response. The method comprises administering an immunostimulatory nucleic

acid to a non-rodent subject in sufficient quantity to stimulate an

immune response. The present sequence is one such immunostimulatory

nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich

(py-rich) or thymidine (T) rich. The method is used to vaccinate subjects

against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae

and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,

haemophilus, campylobacter, clostridium, Escherichia coli and/or

staphylococcus), fungal antigens and/or parasitic antigens. The method is

also useful for preventing cancer, asthma, infectious disease, allergy or

immune deficiency. The present sequence can also be used to redirect a

Th2 to a Th1 immune response and to activate immune cells.

Note: the present sequence may have a phosphorothioate backbone.

XX Sequence 20 BP; 2 A; 3 C; 13 G; 2 T; 0 other;

Query Match 80.0%; Score 16.8; DB 22; Length 20;

Best Local Similarity 90.0%; Pred. No. 3.9e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Search completed: August 10, 2002, 03:24:28
Job time: 13839 sec

```
QY 1 ggggacgacgtcggtggggg 20
    ||||| || |||||
Db 1 ggggacgacgtcggtggggg 20
    ||||| || |||||

RESULT 15
AAF99868
ID AAF99868 standard; DNA; 20 BP.
XX
AC AAF99868;
XX
DT 12-JUN-2001 (first entry)
XX
DE Immunostimulatory nucleic acid #984.
XX
KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS Synthetic.
XX
PN WC200122972-A2.
XX
PD 05-APR-2001.
XX
PF 25-SEP-2000; 2000WO-US26383.
XX
PR 25-SEP-1999; 99US-0156113.
PR 27-SEP-1999; 99US-0156135.
PR 23-AUG-2000; 2000US-0227436.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX
DR WPI; 2001-273485/28.
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma;
PT using immunostimulatory Py-rich and YG nucleic acids -
XX
PS Claim 101; Page 59; 338pp; English.
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SQ Sequence 20 BP; 0 A; 3 C; 13 G; 4 T; 0 other;

Query Match 80.0%; Score 16.8; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 3.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ggggacgacgtcggtggggg 21
    ||||| || |||||
Db 1 ggggacgacgtcggtggggg 20
    ||||| || |||||
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 02:11:29 ; Search time 9068.22 seconds
(without alignments)
31.256 Million cell updates/sec

Title: US-09-672-126-37
Perfect score: 21
Sequence: 1 ggggacgacgtcggtgggggg 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
c 1	17.8	84.8	751	12	AZ174598	AZ174598 SP_0130.A
c 2	17.8	84.8	867	10	BI956859	BI956859 HVSME000
c 3	17.8	84.8	899	12	AZ922331	AZ922331 MRCot5B02
c 4	17.8	84.8	999	12	AG175639	AG175639 Pan trogl
c 5	17.8	84.8	1817	10	BI407913	BI407913 602919349
c 6	17.4	82.9	543	10	BE361794	BE361794 DGL-82.D0
c 7	17.4	82.9	634	10	BE361839	BE361839 DGL-82.D0
c 8	17.4	82.9	657	12	AZ570193	AZ570193 271PVE09
c 9	17.4	82.9	663	12	AZ569270	AZ569270 258PVD06
c 10	17.1	81.0	446	10	BE419896	BE419896 WWS018.G3
c 11	16.8	80.0	148	10	BE595080	BE595080 P11_45.D0
c 12	16.8	80.0	219	10	C90974	C90974 C90974 dict
c 13	16.8	80.0	243	9	AW679219	AW679219 WSL_23.D1
c 14	16.8	80.0	245	9	BS206914	BS206914 BB206914
c 15	16.8	80.0	279	12	BH232306	BH232306 1006167A0
c 16	16.8	80.0	368	10	BE593452	BE593452 WSL_98.A0
c 17	16.8	80.0	379	10	BI779194	BI779194 EBR001_SQ

18	16.8	80.0	384	10	BI098804	BI098804 IPL_33.D1
c 19	16.8	80.0	386	9	AV940016	AV940016 AV940016
c 20	16.8	80.0	397	10	BG834856	BG834856 353220.MA
c 21	16.8	80.0	409	10	BG489315	BG489315 602518437
c 22	16.8	80.0	424	10	BG552821	BG552821 da880c03.
c 23	16.8	80.0	435	9	AW676854	AW676854 DGL_1.C06
c 24	16.8	80.0	448	10	BI098194	BI098194 IPL_29.C0
c 25	16.8	80.0	459	10	BM318700	BM318700 P11_16.F0
c 26	16.8	80.0	467	10	BM323953	BM323953 PIC1_29.F
c 27	16.8	80.0	474	10	BE490177	BE490177 WHE0366.C
c 28	16.8	80.0	486	10	BF624172	BF624172 HVSME001
c 29	16.8	80.0	501	6	BE364338	BE364338 P11_13.F0
c 30	16.8	80.0	517	10	BG053838	BG053838 RH122_10
c 31	16.8	80.0	520	9	AV937915	AV937915 AV937915
c 32	16.8	80.0	541	9	AV939173	AV939173 AV939173
c 33	16.8	80.0	561	9	AV939173	AV939173 AV939173
c 34	16.8	80.0	569	9	AW680468	AW680468 WSL_5.H10
c 35	16.8	80.0	570	9	AV939225	AV939225 AV939225
c 36	16.8	80.0	571	10	BM323482	BM323482 PIC1_19.G
c 37	16.8	80.0	572	9	AV925683	AV925683 AV925683
c 38	16.8	80.0	573	9	AI918263	AI918263 tn09d12.x
c 39	16.8	80.0	597	10	BM324641	BM324641 PIC1_33.G
c 40	16.8	80.0	606	10	BM326628	BM326628 PIC1_59.F
c 41	16.8	80.0	609	9	AV942080	AV942080 AV942080
c 42	16.8	80.0	615	10	BM323410	BM323410 PIC1_18.H
c 43	16.8	80.0	633	10	BE360071	BE360071 DGL_61.F1
c 44	16.8	80.0	690	9	AL508168	AL508168 AL508168
c 45	16.8	80.0	719	9	BE060434	BE060434 HVSMEg001

ALIGNMENTS

RESULT 1
AZ174598/c
LOCUS
DEFINITION
SP_0130.A2_B06_SP6 Strongylocentrotus purpuratus, purple sea urchin
clone Plate=130 Col=12 Row=C, DNA sequence.
511 bp DNA linear GSS 29-AUG-2000
SP_0130.A2_B06_SP6 Strongylocentrotus purpuratus, purple sea urchin
clone Plate=130 Col=12 Row=C, DNA sequence.
AZ174598 GI:8344966
GSS.
Strongylocentrotus purpuratus.
Strongylocentrotus purpuratus.
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 751)
Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,
Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray
,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H. and
Hood,L.
A sea urchin genome project: Sequence scan, virtual map, and
additional resources
Proc. Natl. Acad. Sci. U. S. A. 97 (17), 9514-9518 (2000)
20402566
Contact: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 130 row: C column: 12
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 751.
Location/Qualifiers
1..751
/organism="Strongylocentrotus purpuratus"
/db_xref="taxon:7668"
/clone="Plate=130 Col=12 Row=C"
/clone_lib="Strongylocentrotus purpuratus, purple sea

urchin, sperm genomic BAC library"
/note="Organ: sperm; Vector: BACe3.6; BAC Clones in E-Coli
DH10B"

BASE COUNT 190 a 167 c 125 g 218 t 51 others
ORIGIN

Query Match 84.8%; Score 17.8; DB 12; Length 751;
Best Local Similarity 90.5%; Pred. No. 6.6e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gggagacgactcgtggggggg 21
||||| ||| ||||| |||||
Db 566 GGGGAGGACCTCGTGGGGGG 546

RESULT 2

BI956859

LOCUS HVSMEN0005L01f Hordeum vulgare rachis EST library HVCNDA0015
DEFINITION (normal) Hordeum vulgare cDNA clone HVSMEN0005L01f, mRNA sequence.

ACCESSION BI956859

VERSION BI956859.1 GI:16308112
KEYWORDS EST.

SOURCE barley.

ORGANISM

Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.

REFERENCE

AUTHORS

1 (bases 1 to 867)
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Chin,A., Begum,D.,
Frisch,D., Atkins,M., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons
J., Oates,R. and Main,D.

TITLE

Development of a genetically and physically anchored EST resource
for barley genomics: Morex rachis cDNA library

JOURNAL

COMMENT

Unpublished (2001)
Contact: Wing RA
Clemson University
Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288
Fax: 864 656 4293

Email: rwing@clemson.edu
Total hg bases = 313

Seq primer: AATTAACCTCCTACTAAAGG
High quality sequence stop: 578.

Location/Qualifiers
1. .867

/organism="Hordeum vulgare"
/cultivar="Morex"

/db_xref="taxon:4513"
/clone="HVSMEN0005L01f"

/clone_lib="Hordeum vulgare rachis EST library HVCNDA0015
(normal)"

/tissue_type="Rachis"
/lab_host="JJC121"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XbaI; Plants were grown at Washington State University,
Pullman, WA in a greenhouse, the rachises were excised and
frozen in liquid nitrogen (Kleinhofs lab). In the TJ Close
lab at the University of California, Riverside total RNA
was prepared, poly(A) was purified, one primary
unamplified cDNA library was made, and 1 million pfu were
in vivo excised to give pBluescript SK(-) cDNA phagemids
(Chin). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis see

FEATURES

source

FEATURES

source

1. .899
/organism="Sorghum bicolor"
/cultivar="BTx623"

/db_xref="taxon:4558"
/clone_lib="Sorghum bicolor MRCot"

/tissue_type="leaves"
/dev_stage="seedling"

/note="Vector: pGEM-TA-Easy; A Cot analysis was performed
for the sorghum genome. Based on the resulting Cot curve,
hydroxyapatite chromatography was used to isolate
'highly-repetitive' (HR), 'moderately-repetitive' (MR),
and 'single/low-copy' (SL) sequence components from
sheared genomic DNA. The three repetition-based DNA
components were cloned into E. coli to produce MRCot,
MRCot, and SLcot genomic libraries. Blotting and
sequencing data indicates that each library is
representative of the component from which it was derived.
Putative ID listings given for sequences are based on
comparison (blastn) with sequences in the NCBI Nr
Database. Only the primary match is given (all primary E
values are < or =3D 1.00E-5). In no instance does a 'Cot
clone' contain the complete sequence of its putative Nr
match."

103 a 479 c 215 g 102 t
BASE COUNT

ORIGIN

http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/Bgn/31/cover.html)"

BASE COUNT 157 a 239 c 339 g 130 t 2 others

ORIGIN

Query Match 84.8%; Score 17.8; DB 10; Length 867;
Best Local Similarity 90.5%; Pred. No. 6.6e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gggagacgactcgtggggggg 21
||||| ||| ||||| |||||
Db 664 GGGGAGGCGCGTGGGGGG 684

RESULT 3

AZ922331/c

LOCUS

DEFINITION

AZ922331
MRCot5B02 Sorghum bicolor MRCot Sorghum bicolor genomic, DNA
sequence.

ACCESSION AZ922331

VERSION

AZ922331.1 GI:13432552
KEYWORDS GSS

SOURCE

ORGANISM

Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 899)
Peterson,D.G., Schulze,S.R., Lee,S.A., Sciara,E.B., Nagel,A.,
Tibbitts,D.C., Wessier,S.R. and Paterson,A.H.

Characterization of the Sorghum bicolor genome using DNA
renaturation kinetics (Cot analysis) and repetition-based cloning
Unpublished (2001)
Contact: Peterson DG
Plant Genome Mapping Laboratory
University of Georgia
Room 162, Riverbend Research Bldg., 110 Riverbend Rd., Athens, GA
30602, USA
Tel: 706-583-0167
Fax: 706-583-0160
Email: dgp@arches.uga.edu
Class: Hydroxyapatite-fractionated DNA.
Location/Qualifiers
1. .899

/organism="Sorghum bicolor"
/cultivar="BTx623"

/db_xref="taxon:4558"
/clone_lib="Sorghum bicolor MRCot"

/tissue_type="leaves"
/dev_stage="seedling"

/note="Vector: pGEM-TA-Easy; A Cot analysis was performed
for the sorghum genome. Based on the resulting Cot curve,
hydroxyapatite chromatography was used to isolate
'highly-repetitive' (HR), 'moderately-repetitive' (MR),
and 'single/low-copy' (SL) sequence components from
sheared genomic DNA. The three repetition-based DNA
components were cloned into E. coli to produce MRCot,
MRCot, and SLcot genomic libraries. Blotting and
sequencing data indicates that each library is
representative of the component from which it was derived.
Putative ID listings given for sequences are based on
comparison (blastn) with sequences in the NCBI Nr
Database. Only the primary match is given (all primary E
values are < or =3D 1.00E-5). In no instance does a 'Cot
clone' contain the complete sequence of its putative Nr
match."

103 a 479 c 215 g 102 t
BASE COUNT

ORIGIN

```

Query Match      84.8%; Score 17.8; DB 12; Length 899;
Best Local Similarity 90.5%; Pred. No. 6.7e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgacgtcgtggtggggg 21
||||| || ||||| ||||| |||||
Db 669 GGGGGCGTCTGCTGGGGGG 649

RESULT 4
AG175639/c
LOCUS
DEFINITION Pan troglodytes DNA, clone: RP43-046K06.TJ, genomic survey
sequence.
ACCESSION AG175639
VERSION AG175639.1 GI:16705319
KEYWORDS GSS; GSS (genome survey sequence).
SOURCE Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
Male BAC Library clone:RP43-046K06.TJ.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (sites)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library RPCI-43
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 999)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suhei-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbores@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : ECORI
R.Site 2 : ECORI.
Location/Qualifiers
1..999
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-046K06.TJ"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT 186 a 310 c 133 g 363 t 7 others
ORIGIN

Query Match      84.8%; Score 17.8; DB 12; Length 999;
Best Local Similarity 90.5%; Pred. No. 6.7e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgacgtcgtggtggggg 21
||||| || ||||| ||||| |||||
Db 600 GGGGAAGAGTCTGCTGGGGGG 580

RESULT 5
BI407913/c
LOCUS
DEFINITION BI407913 1817 bp mRNA linear EST 14-AUG-2001
602919349f1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5056045 5',
mRNA sequence.

```

```

ACCESSION BI407913
VERSION BI407913.1 GI:15168836
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1817)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
cDNA Library Preparation: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1152 row: b column: 14
High quality sequence start: 31
High quality sequence stop: 178.
Location/Qualifiers
1..1817
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5056045"
/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTTACCAACTCTGAAGTGGAGCGCGCTGTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 419 a 838 c 285 g 275 t
ORIGIN

Query Match      84.8%; Score 17.8; DB 10; Length 1817;
Best Local Similarity 90.5%; Pred. No. 6.9e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgacgtcgtggtggggg 21
||||| || ||||| ||||| |||||
Db 1750 GCGGAGGACGTCTGTTGGGGGG 1730

RESULT 6
BE361794
LOCUS
DEFINITION BE361794 543 bp mRNA linear EST 20-JUL-2000
DGI_82_D01.bl_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION BE361794
VERSION BE361794.1 GI:9303351
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 543)
AUTHORS Cordonnier-Pratt-M.M., Gingle,A., Marsala,C., Sudman,M. and Pratt
,L.H.
TITLE An EST database from Sorghum: dark-grown seedlings

```

JOURNAL
COMMENT

Unpublished (2000)
Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 537
POLYA-No.

FEATURES

source

Location/Qualifiers

1..543
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGL)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."
91 a 200 c 177 g 75 t

BASE COUNT

ORIGIN

Query Match 82.9%; Score 17.4; DB 10; Length 543;
Best Local Similarity 94.7%; Pred. No. 8.9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 ggacgacgtcgtggggggg 21
|||||
Db 479 GGACGACGTCGTGGCGGG 497

RESULT 7

BE361839

LOCUS

DEFINITION DGL_82_D01.g1_A002 Dark Grown 1 (DGL) Sorghum bicolor cDNA, mRNA
sequence. 634 bp mRNA linear EST 20-JUL-2000

ACCESSION

BE361839

VERSION

BE361839.1

KEYWORDS

EST.

SOURCE

Sorghum bicolor

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE

AUTHORS

Cordonnier-Pratt M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt
L.H.

TITLE

An EST database from Sorghum: dark-grown seedlings

JOURNAL

COMMENT

Unpublished (2000)
Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: PolyTMix
High quality sequence start: 8
High quality sequence stop: 634
POLYA-No.

FEATURES

source

Location/Qualifiers

1..634
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGL)"

/note="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."
BASE COUNT 119 a 181 c 197 g 136 t 1 others
ORIGIN

Query Match 82.9%; Score 17.4; DB 10; Length 634;
Best Local Similarity 94.7%; Pred. No. 9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ggacgacgtcgtggggggg 21

|||||

Db 2 GGACGACGTCGTGGCGGG 20

RESULT 8

AZ570193/c

LOCUS

DEFINITION 271pVE09 Pv MBN #30 Plasmodium vivax genomic 3', DNA sequence. 657 bp DNA linear GSS 15-MAY-2001

ACCESSION

AZ570193

VERSION

AZ570193.1

KEYWORDS

GSS.

SOURCE

malaria parasite P. vivax.

ORGANISM

Plasmodium vivax

REFERENCE

1 (bases 1 to 657)

AUTHORS

Carlton, J.M.-R. and Dame, J.B.

TITLE

The Plasmodium vivax and P. berghei gene sequence tag projects

JOURNAL

COMMENT

Parasitol. Today 16 (10), 409 (2000)

Contact: Dame JB

Dept. of Pathobiology, College of Veterinary Medicine

University of Florida

2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA

Tel: 352 392 4700

Fax: 352 392 9704

Email: damej@mail.vetmed.ufl.edu

Seq primer: M13(-20) forward

Class: shotgun.

FEATURES

Location/Qualifiers

1..657

/organism="Plasmodium vivax"

/strain="Salvador I (Collins, W. 1972. J. Parasitol. 69,

497-598)"

/db_xref="taxon:5855"

/clone_lib="Pv MBN #30"

/dev_stage="asexual blood forms"

/lab_host="Salimiri boliviensis"

/note="Vector: pBluescript SK(+) vector DNA, phagemid
excised from lambda ZAP; Site_1: EcoR V; Site_2: EcoR V;
Host leukocytes were extracted from P. vivax infected
blood using the following methods: first, infected blood
was activated by the addition of 0.5 ml of ADP (40mg/ml)
per 10 ml blood. Then blood was passed over a column of
acid washed 0.1 mm glass beads, then through a Plasmidipur
filter, followed by passage through a column of pre-wet
Whatman CF11 powder (1:2 ratio volume of blood to CF11),
and finally centrifuged through a 50% Percoll density
cushion. Purified DNA was digested with mung bean nuclease
in the presence of 44% formamide at 50°C as described
(Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988.
Nucleic Acids Research 16:6883-6896). Digested DNA was
blunt-ended using T4 DNA Polymerase and size fractionated
over a Sepharose CL-2B column. Fractions in the size range
500bp-4kb were ligated into the Eco RV site of pBluescript
SK(+), and E. coli XL-10 Gold transformed with the
ligation mixture."

BASE COUNT

ORIGIN

146 a 205 c 175 g 129 t 2 others

Query Match

82.9%; Score 17.4; DB 12; Length 657;

Best Local Similarity 94.7%; Pred. No. 9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gggagacgtcgtggggg 20
||||||| |||||||
Db 221 GGGACACATCGTGGGGG 203

RESULT 9
LOCUS AZ569270/c 663 bp DNA linear GSS 15-MAY-2001
DEFINITION 258pVd06 Pv MBN #30 Plasmodium vivax genomic 3', DNA sequence.
ACCESSION AZ569270
VERSION AZ569270.1 GI:13979197
KEYWORDS GSS
SOURCE malaria parasite P. vivax.
ORGANISM Plasmodium vivax
REFERENCE 1 (bases 1 to 663)
AUTHORS The Plasmodium vivax and P. berghei gene sequence tag projects
TITLE Carleton, J.M.-R. and Dame, J.B.
JOURNAL Parasitol. Today 16 (10), 409 (2000)
COMMENT Contact: Dame JB
Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.vetmed.ufl.edu
Seq primer: M13(-20) forward
Class: shotgun.

FEATURES
source
1. 663
/organism="Plasmodium vivax"
/strain="Salvador I (Collins, W. 1972. J. Parasitol. 69, 497-598)."
/db_xref="taxon:5855"
/clone_lib="Pv MBN #30"
/dev_stage="asexual blood forms"
/lab_host="Saimiri boliviensis"
/note="Vector: pBluescript SK(+) vector DNA, phagemid excised from lambda ZAP; Site_1: EcoR V; Site_2: EcoR V. Host leukocytes were extracted from P. vivax infected blood using the following methods: first, infected blood was activated by the addition of 0.5 ml of ADP (40mg/ml) per 10 ml blood. Then blood was passed over a column of acid washed 0.1 mm glass beads, then through a Plasmidipur filter, followed by passage through a column of pre-wet Whatman CF11 powder (1:2 ratio volume of blood to CF11), and finally centrifuged through a 50% Percoll density cushion. Purified DNA was digested with mung bean nuclease in the presence of 4% formamide at 50°C as described (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988. Nucleic Acids Research 16:6883-6896). Digested DNA was blunt-ended using T4 DNA Polymerase and size fractionated over a Sepharose CL-2B column. Fractions in the size range 500bp-4kb were ligated into the Eco RV site of pBluescript SK(+), and E. coli XL-10 Gold transformed with the ligation mixture."

BASE COUNT 144 a 206 c 177 g 133 t 3 others
ORIGIN

Query Match 82.9%; Score 17.4; DB 12; Length 663;
Best Local Similarity 94.7%; Pred. No. 9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gggagacgtcgtggggg 20
||||||| |||||||
Db 234 GGGACACATCGTGGGGG 216

RESULT 10

BE419896
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

FEATURES
source

BASE COUNT
ORIGIN

Query Match 81.0%; Score 17; DB 10; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 gggagacgtcgtgggg 18
||||||| |||||||
Db 93 GGGACACGTCTGGGGG 109

RESULT 11
BE595080
LOCUS

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

BE419896 446 bp mRNA linear EST 24-JUL-2000
WMS018.G3R000101 ITEC WMS Wheat Scutellum Library Triticum aestivum
CDNA clone WMS018.G3, mRNA sequence.

BE419896
BE419896.1 GI:9417742
EST.
bread wheat.

Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 446)

Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier
S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P.,
Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P.,
Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y.,
Pecchioni, N., Qualset, C., Schuch, W., Selvaraj, G., Shariflou, M.,
Sorrenti, M., Warburton, M. and Wenzel, G.
International Triticace EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae
Unpublished (2000)

Contact: Schuch W
zeneca Wheat Improvement Centre, Norwich Research Park
Colney Lane, Norwich NR4 7UH UNITED KINGDOM
Tel: 44 1603 250 2600
Fax: 44 1603 250 699
Email: wolfgang.schuch@aguk.zeneca.com
International Triticace EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.

Location/Qualifiers
1. 446
/organism="Triticum aestivum"
/cultivar="Novosibirskaya 67"
/db_xref="taxon:4565"
/clone="WMS018.G3"
/clone_lib="ITEC WMS Wheat Scutellum Library"
/tissue_type="scutellum callus"
/note="M13 Reverse sequencing primer used for 5' end of clone."

BASE COUNT 132 a 103 c 125 g 86 t

BE595080 148 bp mRNA linear EST 18-AUG-2000
PIL45_D08.bl_A002 Pathogen induced 1 (PIL) Sorghum bicolor cDNA,
mRNA sequence.

BE595080
BE595080.1 GI:9850153
EST.
sorghum.

Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 148)

Cordonnier-Pratt, M.-M., Gingle, A., Dean, R., Sudman, M. and Pratt
L.H.
An EST database from Sorghum: pathogen-induced plants
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Department of Botany

The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.

Seq primer: JEN REV
High quality sequence stop: 41
POLYA-No.

FEATURES 1

source
1. .148
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Pathogen induced 1 (P11)"
/note="Organ: Anthracnose-infected leaves from
two-week-old sorghum plants 48 hr after inoculation;
Vector: pBluescript II from Lambda Zap II; Site_1: XhoI;
Site_2: EcoRI; Two-week-old sorghum plants (BTX 623
cultivar) were infected with pathogen (isolate FRM421 of
Colletotrichum graminicola, which is a sorghum isolate).
RNA was prepared from infected leaves harvested from 45
seedlings 48 hours after inoculation. Note: young
seedlings (2 weeks old) exhibit juvenile resistant
reaction, which is an incompatible interaction. As they
grow older (4 weeks or older), plants resume susceptibility
to anthracnose disease. The library was made from poly-A
RNA in the cloning vector lambda Zap II. Clones to be
sequenced were prepared by mass excision. WARNING: While
most or all ESTs are expected to derive from the host
plant, no effort was made to eliminate ESTs deriving from
the pathogen." 33 a 44 c 48 g 23 t

BASE COUNT

ORIGIN
Query Match 80.0%; Score 16.8; DB 10; Length 148;
Best Local Similarity 90.0%; Pred. No. 1.3e+04;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gggacgacgtcgtggggggg 21

||||| ||||| ||||| |||||

DB 74 GGGACGACGTGTGGCGGGG 93

RESULT 12

C90974/c
LOCUS C90974 219 bp mRNA linear EST 01-MAY-1998
DEFINITION C90974 Dictyostellium discoideum SS (M.Yoshida) Dictyostellium
discoideum cDNA clone SSJ370, mRNA sequence.

ACCESSION C90974

VERSION C90974.1 GI:3097729

KEYWORDS EST.

SOURCE Dictyostellium discoideum.

ORGANISM Dictyostellium discoideum

REFERENCE 1 (bases 1 to 219)

Eukaryota; Mycetozoa; Dictyosteliida; Dictyostellium.

Yoshida, M.

AUTHORS Yoshida, M.

TITLE Developmental cDNA in Dictyostellium discoideum (M.Yoshida)

JOURNAL Unpublished (1998)

COMMENT Contact: Motonobu Yoshida

Research Institute of Food Science

Kinki University

Nakamachi 3327, Nara 631, Japan

Email: yoshida@ews06.nara.kindai.ac.jp.

Location/Qualifiers

1. .219

/organism="Dictyostellium discoideum"

/strain="AX4"

/db_xref="taxon:44689"

/clone="SSJ370"

/clone_lib="Dictyostellium discoideum SS (M.Yoshida)"

BASE COUNT 78 a 69 c 51 g 17 t 4 others
ORIGIN /dev_stage="slug"

Query Match 80.0%; Score 16.8; DB 10; Length 219;

Best Local Similarity 90.0%; Pred. No. 1.4e+04;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gggacgacgtcgtggggggg 21

||||| ||||| ||||| |||||

DB 30 GGGTCGACCTCGTGGGGGG 11

RESULT 13

AW679219

LOCUS AW679219

DEFINITION WSI_23_D10.b1_A002 Water-stressed 1 (WS1) Sorghum bicolor cDNA,

mRNA sequence.

ACCESSION AW679219

VERSION AW679219.1

KEYWORDS EST.

SOURCE sorghum.

ORGANISM Sorghum bicolor

REFERENCE 1 (bases 1 to 243)

AUTHORS Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt

, L.H.

TITLE An EST database from Sorghum: water-stressed plants

JOURNAL Unpublished (2000)

COMMENT Contact: Cordonnier-pratt MM

Department of Botany

The University of Georgia

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 542 1805

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions

below Phred quality 16. The threshold for highest quality sequence

is 20.

Seq primer: JEN REV

High quality sequence stop: 234

POLYA-No.

FEATURES

Location/Qualifiers

1. .243

/organism="Sorghum bicolor"

/db_xref="taxon:4558"

/clone_lib="Water-stressed 1 (WS1)"

/note="Organ: Mix of 5-week old plants on days 7 & 8 after

water was withheld; Vector: Lambda Zap; Site_1: XhoI;

Site_2: EcoRI; The library was made from poly-A RNA in the

cloning vector lambda Zap II. Clones to be sequenced were

prepared by mass excision."

BASE COUNT 53 a 78 c 80 g 32 t

ORIGIN

QY 2 gggacgacgtcgtggggggg 21

||||| ||||| ||||| |||||

DB 89 GGGACGACGTGTGGCGGGG 108

RESULT 14

BB206914/c

LOCUS BB206914

DEFINITION BB206914 RIKEN full-length enriched, 0 day neonate thymus Mus

musculus cDNA clone A430078M12 3' similar to U39827 Mus musculus

musculus cDNA clone A430078M12 3' similar to U39827 Mus musculus

putative G protein-coupled receptor TDAG8 (TDAG8) mRNA, mRNA
sequence.
BB206914
BB206914.1 GI:8871867
EST.
house mouse.
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 245)
Konno.H., Aizawa.K., Akahira.S., Akiyama.J., Arahawa.T., Carninci
P., Endo.T., Fukuda.S., Fukunishi.Y., Hara.A., Hayatsu.N.,
Hirozane.T., Hori.F., Ishii.Y., Ishikawa.J., Ishikawa.T., Itoh.M.,
Izawa.M., Kadota.K., Kagawa.I., Kai.C., Kawai.J., Kikuchi.N.,
Kiyosawa.H., Kojima.Y., Kondo.S., Koya.S., Kurihara.C., Kusakabe.M.,
Matsuyama.T., Miki.R., Mizuno.Y., Nakamura.M., Oda.H., Okazaki.Y.,
Ono.T., Owa.C., Saito.H., Sakai.C., Sato.K., Shibata.K., Shibata
Y., Shigemoto.Y., Shinagawa.A., Shiraki.T., Sogabe.Y., Sugahara.Y.,
Suzuki.H., Suzuki.H., Tagawa.A., Takahashi.F., Tomimaga.N., Toya
T., Tsunoda.Y., Watahiki.A., Watanabe.S., Yamamura.T., Yamanaka.I.,
Yano.R., Yasunishi.A., Yokota.T., Yoshida.K., Yoshiki.A., Yoshino
M., Muramatsu.M. and Hayashizaki.Y.
RIKEN Mouse ESTs (Konno.H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci.P., Nishiyama.Y., Westover.A., Itoh.M., Nagaoka.S., Sasaki
N., Okazaki.Y., Muramatsu.M. and Hayashizaki.Y.
Thermotabilization and thermoactivation of thermostable enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh.M., Kitsuai.T., Akiyama.J., Shibata.K., Izawa.M., Kawai.J.,
Tomaru.Y., Carninci.P., Shibata.Y., Ozawa.Y., Muramatsu.M., Okazaki
Y. and Hayashizaki.Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci.P. and Hayashizaki.Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.

FEATURES

source

Location/Qualifiers
1..245
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A430078M12"
/clone_lib="RIKEN full-length enriched, 0 day neonate
thymus"
/tissue_type="thymus"
/dev_stage="0 day neonate"
/lab_host="DH10B"

/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCGCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adaptor of
sequence [5' GAGAGAGAGATCCAGTAAATTAATATCCCTCCCTCCCTCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a

modified pBluescript KS(+) after bulk excision from Lambda
FLC I."
BASE COUNT 49 a 85 c 39 g 62 t 10 others
ORIGIN

Query Match 80.0%; Score 16.8; DB 9; Length 245;
Best Local Similarity 85.7%; Pred. No. 1.4e+04;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggggacgacgtcgtggggggg 21
||||||| | |||||
Db 42 GGGGACGANGGGTGGGGGG 22

RESULT 15
BH232306/c
LOCUS
DEFINITION 1006167A02.y1 1006 - RescueMu Grid G Zea mays genomic, DNA
ACCESSION BH232306
VERSION BH232306.1 GI:16837376
KEYWORDS GSS.
SOURCE
ORGANISM Zea mays
ze mays

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 279)
Walbot.V.
Maize genomic sequences found using engineered RescueMu transposon
Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1006167 row: 23
Class: transposon-tagged.
Location/Qualifiers
1..279
/organism="Zea mays"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/clone_lib="1006 - RescueMu Grid G"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"

/note="Organ: leaf; Vector: RescueMu (engineered from
pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
extracted from leaf punches, double digested using BamHI
and BglII, and ligated to form circular plasmids. DH10B
cells were transformed and then screened on LB plates with
ampicillin."

BASE COUNT 35 a 121 c 56 g 67 t
ORIGIN

Query Match 80.0%; Score 16.8; DB 12; Length 279;
Best Local Similarity 90.0%; Pred. No. 1.4e+04;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ggggacgacgtcgtggggggg 21

Db 96 GCGAGGACGTCGTGGGGGG 77

Search completed: August 10, 2002, 02:11:32
Job time: 13153 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 03:09:20 ; Search time 277.54 Seconds
(without alignments)
18.586 Million cell updates/sec

Title: US-09-672-126-37
Perfect score: 21
Sequence: 1 ggggacgactgctggggggg 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	16.2	77.1	1053	6	Patent No. 5352575
C 2	16.2	77.1	1386	1	Sequence 4, Appli
C 3	16.2	77.1	1473	1	Sequence 2, Appli
C 4	15.8	75.2	4190	3	Sequence 2, Appli
C 5	15.8	75.2	4403765	4	Sequence 2, Appli
C 6	15.4	73.3	4522	5	Sequence 22, Appli
C 7	15.2	72.4	143	4	Sequence 263, App
C 8	15.2	72.4	234	1	Sequence 12, Appl
C 9	15.2	72.4	234	1	Sequence 12, Appl
C 10	15.2	72.4	234	5	Sequence 12, Appl
C 11	15.2	72.4	266	4	Sequence 2, Appli
C 12	15.2	72.4	456	2	Sequence 16, Appl
C 13	15.2	72.4	456	3	Sequence 16, Appl
C 14	15.2	72.4	456	4	Sequence 16, Appl
C 15	15.2	72.4	701	4	Sequence 1, Appli
C 16	15.2	72.4	765	4	Sequence 29, Appl
C 17	15.2	72.4	816	4	Sequence 31, Appl
C 18	15.2	72.4	894	1	Sequence 10, Appl
C 19	15.2	72.4	894	1	Sequence 3, Appli
C 20	15.2	72.4	894	2	Sequence 3, Appli
C 21	15.2	72.4	894	2	Sequence 3, Appli
C 22	15.2	72.4	894	4	Sequence 3, Appli
C 23	15.2	72.4	1008	1	Sequence 8, Appli
C 24	15.2	72.4	1008	1	Sequence 1, Appli
C 25	15.2	72.4	1008	1	Sequence 1, Appli
C 26	15.2	72.4	1008	1	Sequence 1, Appli
C 27	15.2	72.4	1008	1	Sequence 1, Appli

C 28	15.2	72.4	1008	1	US-08-485-978-1	Sequence 1, Appli
C 29	15.2	72.4	1008	2	US-08-481-970-1	Sequence 1, Appli
C 30	15.2	72.4	1008	2	US-08-486-814-1	Sequence 1, Appli
C 31	15.2	72.4	1008	2	US-08-487-472-1	Sequence 1, Appli
C 32	15.2	72.4	1008	2	US-08-897-719-1	Sequence 1, Appli
C 33	15.2	72.4	1008	3	US-08-485-740-1	Sequence 1, Appli
C 34	15.2	72.4	1008	3	US-09-162-184-1	Sequence 1, Appli
C 35	15.2	72.4	1008	4	US-09-161-902-1	Sequence 1, Appli
C 36	15.2	72.4	1008	4	US-09-163-269-1	Sequence 1, Appli
C 37	15.2	72.4	1008	4	US-09-489-717A-1	Sequence 1, Appli
C 38	15.2	72.4	1008	5	PCT-US95-08179-1	Sequence 1, Appli
C 39	15.2	72.4	1095	4	US-08-983-035A-25	Sequence 25, Appl
C 40	15.2	72.4	1128	4	US-08-983-035A-27	Sequence 27, Appl
C 41	15.2	72.4	1149	4	US-08-983-035A-35	Sequence 35, Appl
C 42	15.2	72.4	1167	4	US-09-411-687A-4	Sequence 4, Appli
C 43	15.2	72.4	1209	4	US-08-983-035A-33	Sequence 33, Appl
C 44	15.2	72.4	1947	4	US-09-025-769B-264	Sequence 264, App
C 45	15.2	72.4	1956	3	US-08-693-940-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

5352575-6/C

Patent No. 5352575

APPLICANT: PETROVSKIS, ERIK A.;POST, LEONARD E.;TIMMINS, JAMES G.

TITLE OF INVENTION: PSEUDORABIES VIRUS PROTEIN

NUMBER OF SEQUENCES: 12

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/513,282

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 100,817

FILING DATE: 29-JUN-1987

APPLICATION NUMBER: 886,260

FILING DATE: 16-JUL-1986

APPLICATION NUMBER: 784,787

FILING DATE: 04-OCT-1985

APPLICATION NUMBER: 801,799

FILING DATE: 26-NOV-1985

APPLICATION NUMBER: 844,113

FILING DATE: 26-MAR-1986

SEQ ID NO:6

LENGTH: 1053

5352575-6

Query Match 77.1%; Score 16.2; DB 6; Length 1053;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ggggacgactgctggggggg 21

Db 536 GCGGACGACGGCGTGGTGG 516

RESULT 2

US-08-672-571A-4

Sequence 4, Application US/08672571A

Patent No. 5795765

GENERAL INFORMATION:

APPLICANT: IZU, Hiroyuki

APPLICANT: KURUME, Yoko

APPLICANT: IZUMI, Yoshiya

APPLICANT: SANO, Mutsumi

APPLICANT: KATO, Ikunoshin

APPLICANT: ITO, Makoto

TITLE OF INVENTION: Gene Encoding Endoglycoceramidase

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP

STREET: P.O. Box 747

;/ CITY: Falls Church
;/ STATE: Virginia
;/ COUNTRY: USA
;/ ZIP: 22040-0747
;/
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/672,571A
;/ FILING DATE: 28 JUNE 1996
;/ CLASSIFICATION: 435
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: WEINER, Marc S.
;/ REGISTRATION NUMBER: 32,181
;/ REFERENCE/DOCKET NUMBER: 1422-0264P
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (703) 205-8000
;/ TELEFAX: (703) 205-8050
;/ TELEX: 248345
;/
;/ INFORMATION FOR SEQ ID NO: 4:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 1386 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: double
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: genomic DNA
;/
;/ US-08-672-571A-4

Query Match 77.1%; Score 16.2; DB 1; Length 1386;
Best Local Similarity 85.7%; Pred. No. 76;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 999gacgcgtcgtggtggggg 21
||| ||||| ||||| |||

Db 1285 GCGCGGACGTCGTGGGTGG 1305

RESULT 3

US-08-672-571A-2
; Sequence 2, Application US/08672571A
; Patent No. 5795765
; GENERAL INFORMATION:
; APPLICANT: IZU, Hiroyuki
; APPLICANT: KURUME, Yoko
; APPLICANT: IZUMI, Yoshiya
; APPLICANT: SANO, Mutsumi
; APPLICANT: KATO, Ikunoshin
; APPLICANT: ITO, Makoto
; TITLE OF INVENTION: Gene Encoding Endoglycoceramidase
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,571A
; FILING DATE: 28 JUNE 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-0264P

;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (703) 205-8000
;/ TELEFAX: (703) 205-8050
;/ TELEX: 248345
;/
;/ INFORMATION FOR SEQ ID NO: 2:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 1473 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: double
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: genomic DNA
;/
;/ US-08-672-571A-2

Query Match 77.1%; Score 16.2; DB 1; Length 1473;
Best Local Similarity 85.7%; Pred. No. 76;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 999gacgcgtcgtggtggggg 21
||| ||||| ||||| |||

Db 1372 GCGCGGACGTCGTGGGTGG 1392

RESULT 4

US-08-938-291A-2/c
; Sequence 2, Application US/08938291A
; Patent No. 6117673
; GENERAL INFORMATION:
; APPLICANT: Lev, Sima
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: RDB PROTEINS AND RELATED
; PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,291A
; FILING DATE: September 26, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,337
; FILING DATE: October 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 228/172
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4190 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;/
;/ US-08-938-291A-2

Query Match 75.2%; Score 15.8; DB 3; Length 4190;

Best Local Similarity 89.5%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgacgtctgtgggg 19
||| ||||| ||||| |||||

Db 3446 GGAGACGACGCGTGGGG 3428

RESULT 5

US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328

GENERAL INFORMATION:

APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

TITLE OF INVENTION: TUBERCULOSIS

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A

CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 4403765

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

FEATURE:

OTHER INFORMATION: CDC 1551

OTHER INFORMATION: "n" bases at various positions throughout the sequence

OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Query Match 75.2%; Score 15.8; DB 4; Length 4403765;
Best Local Similarity 89.5%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggggacgacgtctgtgggg 19
||| ||||| ||||| |||||

Db 4173327 GGTGACGACGTCGTGGG 4173309

RESULT 6

PCT-US93-06251-22/c

Sequence 22, Application PC/TUS9306251

GENERAL INFORMATION:

APPLICANT: Wickstrom, Eric and Rife, Jason P.

TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing

TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: NY

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/06251

FILING DATE: 19930630

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: DiGiglio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 8586

TELECOMMUNICATION INFORMATION:

TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 4522 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-22

Query Match 73.3%; Score 15.4; DB 5; Length 4522;
Best Local Similarity 94.1%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 gacgacgtctgtggggg 20
||||| ||||| |||||

Db 3211 GACGAGTCTGTGGGGG 3195

RESULT 7

US-09-025-769B-263/c

Sequence 263, Application US/09025769B

Patent No. 6300064

GENERAL INFORMATION:

APPLICANT: Knappik, Achim

APPLICANT: Pack, Peter

APPLICANT: Ilag, Vic

APPLICANT: Ge, Liming

APPLICANT: Moroney, Simon

APPLICANT: Plueckthun, Andreas

TITLE OF INVENTION: Protein/(Poly)peptide libraries

NUMBER OF SEQUENCES: 373

CORRESPONDENCE ADDRESS:

ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10021

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/025,769B

FILING DATE: 18-FEB-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 95 11 3021.0

FILING DATE: 18-AUG-1995

ATTORNEY/AGENT INFORMATION:

NAME: James F. Haley, Jr., Esq.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: MORPHO/5

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)596-9000

TELEFAX: (212)596-9090

INFORMATION FOR SEQ ID NO: 263:

SEQUENCE CHARACTERISTICS:

LENGTH: 143 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "synthetic DNA cassette"

US-09-025-769B-263

Query Match 72.4%; Score 15.2; DB 4; Length 143;
Best Local Similarity 85.0%; Pred. No. 2.2e+02;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gggagcagctcgtggggggg 21
||| ||||| |||||
Db 99 GGGGGGAGTCGGGGGGG 80

RESULT 8

US-08-347-792-12/c
; Sequence 12, Application US/08347792

; Patent No. 5573925

; GENERAL INFORMATION:

; APPLICANT: Halazonetis, Thanos D.

; TITLE OF INVENTION: p53 Proteins With Altered

; TITLE OF INVENTION: Tetramerization Domains

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson

; STREET: Spring House Corporate Cntr., PO Box 457

; CITY: Spring House

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/347,792

; FILING DATE:

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Bak, Mary E.

; REGISTRATION NUMBER: 31,215

; REFERENCE/DOCKET NUMBER: WST58USA

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-540-9206

; TELEFAX: 215-540-5818

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 234 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..234

US-08-347-792-12

Query Match 72.4%; Score 15.2; DB 1; Length 234;

Best Local Similarity 85.0%; Pred. No. 2.2e+02;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggagcagctcgtggggggg 20
||| ||| ||||| |||||
Db 156 GGGGGGAGTCGTGGGGG 137

RESULT 9

US-08-431-357-12/c

; Sequence 12, Application US/08431357

; Patent No. 5721340

; GENERAL INFORMATION:

; APPLICANT: Halazonetis, Thanos D.

; TITLE OF INVENTION: p53 Proteins With Altered

; TITLE OF INVENTION: Tetramerization Domains

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson

; STREET: Spring House Corporate Cntr., PO Box 457

; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,357
; FILING DATE:
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/347,792

; FILING DATE: 28-NOV-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Bak, Mary E.

; REGISTRATION NUMBER: 31,215

; REFERENCE/DOCKET NUMBER: WST58USA

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-540-9206

; TELEFAX: 215-540-5818

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 234 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..234

US-08-431-357-12

Query Match 72.4%; Score 15.2; DB 1; Length 234;

Best Local Similarity 85.0%; Pred. No. 2.2e+02;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gggagcagctcgtggggggg 20
||| ||| ||||| |||||
Db 156 GGGGGGAGTCGTGGGGG 137

RESULT 10

PCT-US95-15353-12/c

; Sequence 12, Application PC/TUS9515353

; GENERAL INFORMATION:

; APPLICANT: The Wistar Institute of Anatomy

; APPLICANT: and Biology

; APPLICANT: Halazonetis, Thanos D.

; TITLE OF INVENTION: p53 Proteins With Altered

; TITLE OF INVENTION: Tetramerization Domains

; NUMBER OF SEQUENCES: 46

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson

; STREET: Spring House Corporate Cntr., PO Box 457

; CITY: Spring House

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/15353

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/347,792


```

; LENGTH: 701 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 17..694
US-09-133-321-1

```

```

Query Match      72.4%; Score 15.2; DB 4; Length 701;
Best Local Similarity 85.0%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 ggggacgacgtcgtggggg 20
    ||||| || ||||| |||||
Db 616 GGGGGCGGAGTCGTGGGGG 597

```

```

Search completed: August 10, 2002, 03:10:28
Job time: 16294 sec

```

THIS PAGE BLANK (USPTO)